

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 31.9318 Seconds
(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-8

Perfect score: 25

Sequence: 1 DXNDN 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	96.0	15	16	AA198281
2	24	96.0	20	18	AA198255
3	24	96.0	20	21	AA198895
4	24	96.0	40	15	AA198887
5	24	96.0	40	17	AA1987126
6	24	96.0	43	15	AA1988860
7	24	96.0	43	15	AA1988903
8	24	96.0	43	17	AA1987102
9	24	96.0	43	17	AA1987143
10	24	96.0	45	15	AA1986073

RESULT 1
AAR78281
ID AAR78281 standard; peptide; 15 AA.

ALIGNMENTS

11	24	96.0	60	23	ABP32657	Human cadherin-lik
12	24	96.0	81	21	AA198433	Haematocobla irritan
13	24	96.0	86	22	AB142873	Peptide #10379 enc
14	24	96.0	86	22	AB142873	Protein #10379 enc
15	24	96.0	86	22	AB142873	Human brain expres
16	24	96.0	86	22	AB142873	Human bone marrow
17	24	96.0	86	22	AB142873	Peptide #10727 enc
18	24	96.0	86	22	AA198690	Human novel extrac
19	24	96.0	91	22	AA198683	Human polypeptide
20	24	96.0	91	22	AB147903	Nonclassical cadhe
21	24	96.0	107	21	AA198414	Variant variable h
22	24	96.0	120	17	AA198414	Anti-human FasL an
23	24	96.0	120	18	AA198414	Variable heavy cha
24	24	96.0	121	17	AA198414	Human ORFX protein
25	24	96.0	131	23	AB198216	Human polypeptide
26	24	96.0	134	22	AA198589	Human polypeptide
27	24	96.0	136	23	AB198577	Human ORFX protein
28	24	96.0	140	21	AB198577	Bacteriophage 192
29	24	96.0	142	22	AA198508	Human polypeptide
30	24	96.0	146	15	AA198577	Rat-413 cadherin p
31	24	96.0	146	17	AA198577	Protocadherin clon
32	24	96.0	156	21	AA198577	Human ORFX ORF2113
33	24	96.0	165	22	AA198577	Human novel extrac
34	24	96.0	165	23	AB198577	Human polypeptide
35	24	96.0	168	21	AA198577	Haematocobla irritan
36	24	96.0	175	21	AA198577	Haematocobla irritan
37	24	96.0	177	22	AA198577	G protein-coupled
38	24	96.0	188	17	AA198577	Peptide sequence.
39	24	96.0	188	18	AA198577	Thermococcus prote
40	24	96.0	193	22	AA198577	Gene 15 related pe
41	24	96.0	193	22	AA198577	Human secreted pro
42	24	96.0	205	22	AA198577	Gene 15 related pe
43	24	96.0	205	22	AA198577	C glutamicum prote
44	24	96.0	212	22	AA198577	Hyperthermostable
45	24	96.0	237	17	AA198577	S. epidermidis ope
46	24	96.0	265	22	AA198577	Human secreted pro
47	24	96.0	285	21	AA198577	Bacillus subtilis
48	24	96.0	291	19	AA198577	Drosophila melanog
49	24	96.0	314	22	AB198577	Novel human diagno
50	24	96.0	321	22	AB198577	Drosophila melanog
51	24	96.0	328	22	AB198577	Human secreted pro
52	24	96.0	331	21	AA198577	Staphylococcus epi
53	24	96.0	332	23	AB198577	Novel human diagno
54	24	96.0	334	22	AB198577	Novel human diagno
55	24	96.0	336	22	AB198577	Human novel extrac
56	24	96.0	339	22	AA198577	Human polypeptide
57	24	96.0	339	23	AB198577	Mouse wound healin
58	24	96.0	341	22	AB198577	Human mddt protein
59	24	96.0	355	22	AA198577	Novel human diagno
60	24	96.0	362	22	AB198577	Outer membrane lip
61	24	96.0	367	15	AA198577	Human wound healin
62	24	96.0	370	22	AB198577	Novel human diagno
63	24	96.0	378	22	AB198577	Novel human diagno
64	24	96.0	380	22	AB198577	Human wound healin
65	24	96.0	382	22	AB198577	Novel human diagno
66	24	96.0	388	22	AB198577	Human wound healin
67	24	96.0	390	22	AB198577	Human protein sequ
68	24	96.0	398	22	AB198577	Human protein sequ
69	24	96.0	413	22	AB198577	Drosophila melanog
70	24	96.0	419	22	AB198577	Novel human diagno
71	24	96.0	421	22	AB198577	Novel human diagno
72	24	96.0	427	22	AA198577	Human novel extrac
73	24	96.0	433	23	AB198577	Human polypeptide
74	24	96.0	433	20	AA198577	Salmonella dublin
75	24	96.0	443	22	AA198577	S. epidermidis ope

XX AAR78281;
 XX
 DT 13-MAR-1996 (first entry)
 XX
 DE Yellow fever virus derived peptide.
 XX
 KW Yellow fever virus; positive control.
 XX
 OS Flavivirus febricis.
 XX
 PN CA2137640-A.
 PD 16-JUN-1995.
 XX
 PF 08-DEC-1994; 94CA-2137640.
 XX
 PR 15-DEC-1993; 93US-0167336.
 XX
 PA (HEAL-) HEALTH RES INC.
 PA (UNLO) UNIV COLLEGE LONDON.
 XX
 PI Pride M, Roitt I, Thakur A, Thanavala Y;
 DR WPI; 1995-275760/37.
 XX
 PT Anti-idiotypic antibody 2F10 mimicking hepatitis B surface antigen
 PT - useful in vaccine(s) or to detect successful vaccination or
 PT carrier(s).
 XX
 PS Disclosure; Page 13; 52pp; English.
 XX
 CC AAR78281 is a Yellow fever virus derived peptide, used as a positive
 CC control in antigen processing experiments.
 CC
 SQ Sequence 15 AA;
 Query Match 96.0%; Score 24; DB 16; Length 15;
 Best Local Similarity 80.0%; Pred. No. 75;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 8 DTNDN 12
 RESULT 2
 AAW40255
 ID AAW40255 standard; Protein; 20 AA.
 XX
 AC AAW40255;
 DT 15-JUN-1998 (first entry)
 XX
 DE Human wild-type E-Cadherin fragment (pos. 360-379).
 XX
 KW E-Cadherin; mutant; variant; monoclonal antibody; diagnosis; therapy;
 KW gastric carcinoma; immunoassay; therapeutic conjugate; toxin;
 KW tumour cell.
 XX
 OS Homo sapiens.
 XX
 PN DE19629938-C1.
 PD 27-NOV-1997.
 XX
 PF 24-JUL-1996; 96DE-1029938.
 XX
 PR 24-JUL-1996; 96DE-1029938.
 XX
 PA (GSFU-) GSF-FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
 XX
 PI Becker K, Eulitz M, Hoefler H, Krammer E, Schuhmacher C;

XX WPI; 1997-551542/51.
 XX
 PT Monoclonal antibodies specific for mutated E-Cadherin peptide
 XX sequences - useful for diagnosis and therapy of stomach cancer
 XX
 PS Disclosure; Page 7; 23pp; German.
 XX
 CC This sequence represents a fragment of the E-Cadherin protein which is
 CC used in the construction of the mutant represented in AAW36609. This
 CC fragment is used in a method where a novel monoclonal antibody
 CC recognises one or more mutated (through deletion or point mutation)
 CC E-Cadherin sequences. Such antibodies can be used for diagnosis and
 CC therapy of gastric carcinomas, especially as diagnostic immunoassay
 CC reagents or as therapeutic conjugates with toxins or radionuclides. The
 CC oligonucleotides can be used to detect tumour cells in samples containing
 CC human cells. The oligopeptides can also be used for immunotherapy of
 CC tumours, especially gastric carcinomas.
 XX
 SQ Sequence 20 AA;
 Query Match 96.0%; Score 24; DB 18; Length 20;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 8 DTNDN 12
 RESULT 3
 AAY54895
 ID AAY54895 standard; peptide; 20 AA.
 XX
 AC AAY54895;
 DT 09-FEB-2000 (first entry)
 XX
 DE Wild type E-cadherin fragment.
 XX
 KW E-Cadherin; mutein; monoclonal antibody; detection; diagnosis;
 KW gastric cancer.
 XX
 OS Homo sapiens.
 XX
 PN JPI1292900-A.
 PD 26-OCT-1999.
 XX
 PF 20-JAN-1998; 98JP-0042794.
 XX
 PR 20-JAN-1998; 98JP-0042794.
 XX
 PA (GSFU-) GSF-FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
 XX
 DR WPI; 2000-018716/02.
 XX
 PT New E-cadherin mutant - useful as basis for diagnosis and treatment of
 PT human malignant tumour
 XX
 PS Disclosure; Page 11; 33pp; Japanese.
 XX
 CC This sequence represents a fragment of wild type E-cadherin protein.
 CC The invention relates to a monoclonal antibody which is directed
 CC specifically to the amino acid sequence of a mutated E-cadherin and is
 CC formed by an intraquadrant mutation at DNA level. The mutation at least:
 CC (a) leads to the loss of at least one base triplet or its polymer in an
 CC exon at the RNA level and successively leads to the deletion of at least
 CC one amino acid of the wild type E-cadherin protein; or (b) leads to the
 CC exchange of 1 or 2 nucleotides of at least one base triplet in an exon at
 CC the RNA level and successively leads to the exchange of at least one
 CC amino acid of the wild type E-cadherin protein. The monoclonal antibody is
 CC useful for the detection of gastric cancer. The monoclonal antibody is

CC useful for the specific detection of scattered gastric cancer.

XX Sequence 20 AA;

SQ Query Match 96.0%; Score 24; DB 21; Length 20;

Best Local Similarity 80.0%; Pred. No. 1e+02; Mismatches 0; Indels 1; Gaps 0;

Matches 4; Conservative 0;

Qy 1 DXNDN 5

Db 8 DTNDN 12

RESULT 4

AAR5887

ID AAR5887 standard; Protein; 40 AA.

XX AC AAR5887;

XX DT 17-APR-1995 (first entry)

XX DE Human-13 cadherin-related molecule.

XX KW Cadherin; cell adhesion molecule.

XX OS Homo sapiens.

XX PN WO9414960-A.

XX PD 07-JUL-1994.

XX PF 23-DEC-1993; 93WO-US12588.

XX PR 29-DEC-1992; 92US-0998003.

XX PA (DOHE-) DOHENY EYE INST.

XX PI Suzuki S;

XX DR WPI; 1994-293849/36.

XX DR N-PSDB; AAR58877.

XX PT Polynucleotide sequences encoding new proto:cadherins - useful

XX PS for modulating natural binding and regulating activities.

XX Example; Page 52; 114pp; English.

XX CC Two regions of conserved AA sequence, one from the middle of the

XX CC third cadherin extracellular subdomain (EC-3) and the other from the

XX CC C-terminus of the fourth extracellular subdomain (EC-4) were

XX CC identified. The corresp. degenerate oligos (AAQ68949, AAQ68950) were

XX CC designed for use as PCR primers. PCR was carried out on a rat brain

XX CC cDNA prep. Two major bands of about 450 bps and 130 bps were found.

XX CC The 450 bp band corresponded to the expected length between the two

XX CC primer sites, but the 130 bp band could not be predicted from any

XX CC of the previously identified cadherin sequences. The 450 bp and 130

XX CC bp bands were extracted and sequenced. Nineteen novel partial cDNA

XX CC clones were isolated. The DNA and deduced AA sequences of the

XX CC clones (including sequences corresp. to the PCR primers) are given

XX CC in AAQ68951-068969 and AAR58860-R58878. Various cDNA fragments

XX CC structurally similar to the rat cDNAs were isolated from human,

XX CC mouse and Xenopus brain cDNA preps. and from Drosophila and

XX CC C. elegans whole body cDNA preps. By PCR using the above primers.

XX CC The DNA and deduced AA sequences of the resulting PCR fragments

XX CC (including sequences corresp. to the PCR primers) are given in

XX CC AAQ68971, AAQ68972-068994 and AAR58882-R58905 and AAR49143. Comparison

XX CC of the deduced AA sequences indicates a similarity, in particular,

XX CC there are three sets of clones that appear to be cross-species

XX CC homologues: RAT-218, MOUSE-322 and HUMAN-43; RAT-314, MOUSE-321 and

XX CC HUMAN-11; and MOUSE-326 and HUMAN-42.

XX SQ Sequence 40 AA;

Query Match

Best Local Similarity 96.0%; Score 24; DB 15; Length 40;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5

Db 34 DANDN 38

RESULT 5

AAR87126

ID AAR87126 standard; Peptide; 40 AA.

XX AC AAR87126;

XX DT 29-AUG-1996 (first entry)

XX DE Protocadherin clone HUMAN-13.

XX KW Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion; mouse;

XX OS catenin; therapy; clone.

XX PN Homo sapiens.

XX WO9600289-A1.

XX PD 04-JAN-1996.

XX PF 26-JUN-1995; 95WO-US08071.

XX PR 27-JUN-1994; 94US-0268161.

XX PA (DOHE-) DOHENY EYE INST.

XX PI Suzuki S;

XX DR WPI; 1996-068873/07.

XX DR N-PSDB; AAT03601.

XX PT Polynucleotide(s) encoding human protocadherins pc3 and pc4 and rat

XX PS pc5 - involved in cell-cell adhesion and regulation activities

XX Example 2; Page 55-56; 146pp; English.

XX CC AAR87125-R87137 represent partial fragments of the human protocadherin

XX CC sequence. The cDNAs encoding these sequences were isolated after

XX CC screening a human brain cDNA preparation with the primers shown in

XX CC AAT03575 and AAT03576. The primers were constructed from portions of the

XX CC amino acid sequences of the third and fourth extracellular domains of

XX CC published cadherin sequences. The full length cDNA sequences encoding

XX CC human protocadherins pc3 and pc4 are represented by AAT03572 and

XX CC AAT03573, respectively. The cytoplasmic domain of cadherin interacts

XX CC with the cytoskeleton through catenins and other cytoskeleton associated

XX CC proteins. The cytoplasmic domain is not present in all cadherins, but in

XX CC those which possess it, it is essential for the cadherin adhesive

XX CC function. The cadherins which do not possess a cytoplasmic domain appear

XX CC to function via a different method from those with a cytoplasmic domain.

XX CC These protein sequences are involved in cell-cell adhesion. These

XX CC sequences may have regulatory functions in the cell, as well as the

XX CC cell-cell adhesive properties. Antibodies produced against these

XX CC sequences are useful for modulating the binding activity of these

XX CC protocadherins, and can be used therapeutically.

XX SQ Sequence 40 AA;

Query Match

Best Local Similarity 80.0%; Score 24; DB 17; Length 40;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5

Db 34 DANDN 38

RESULT 6
 AAR58860
 ID AAR58860 standard; Protein; 43 AA.
 XX
 AC AAR58860;
 XX
 DT 17-APR-1995 (first entry)
 XX
 DE Rat-123 cadherin partial sequence.
 XX
 KW Cadherin; cell adhesion molecule.
 XX
 OS Rattus rattus.
 XX
 PN WO9414960-A.
 XX
 PD 07-JUL-1994.
 XX
 PF 23-DEC-1993; 93WO-US12588.
 XX
 PR 29-DEC-1992; 92US-0998003.
 XX
 PA (DOHE-) DOHENY EYE INST.
 XX
 PI Suzuki S;
 XX
 DR WPI; 1994-293849/36.
 DR N-PSDB; AAR58860.
 XX
 PT Polynucleotide sequences encoding new proto:cadherins - useful
 PT for modulating natural binding and regulating activities.
 XX
 PS Example; Page 34; 114pp; English.
 XX
 CC Two regions of conserved AA sequence, one from the middle of the
 CC third cadherin extracellular subdomain (EC-3) and the other from the
 CC C-terminus of the fourth extracellular subdomain (EC-4) were
 CC identified. The corresp. degenerate oligos (Q68949, Q68950) were
 CC designed for use as PCR primers. PCR was carried out on a rat brain
 CC cDNA prep. Two major bands of about 450 bps and 130 bps were found.
 CC The 450 bp band corresponded to the expected length between the two
 CC primer sites, but the 130 bp band could not be predicted from any
 CC of the previously identified cadherin sequences. The 450 bp and 130
 CC bp bands were extracted and sequenced. Nineteen novel partial cDNA
 CC clones (including sequences corresp. to the PCR primers) are given
 CC in Q68951-Q68969 and R58860-R58878. The deduced AA sequences of the
 CC cDNA clones are homologous to, but distinct from the known
 CC cadherins. The cadherins described thus far have highly conserved
 CC short AA sequences in the EC-3 including the consensus sequence
 CC D-Y-E or D-E-F located at the middle region of the subdomain and
 CC the consensus sequence in R58879 or R58880 at its end, while the
 CC corresp. sequences of other subdomains, except for the 5th extra-
 CC cellular subdomain (EC-5), are D-R-E and the sequence in R58881
 CC respectively. In contrast the deduced AA sequences of the new
 CC clones that corresp. to cadherin extracellular subdomains include
 CC the sequence D-Y-E or D-E-F at one end, but have the sequence
 CC D-X-N-D-N-X-P-X-F instead of R58879 or R58880 at the other end.
 CC The polypeptides encoded by the partial clones are homologous to
 CC previously identified cadherins but did not show significant
 CC homology to any other sequences in Genbank. Therefore, the partial
 CC cDNAs appear to comprise a new subclass of cadherin-related
 CC molecules.
 XX
 SQ Sequence 43 AA;
 Query Match 96.0%; Score 24; DB 15; Length 43;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 |
 |
 |
 DB 37 DANDN 41

DB 37 DANDN 41
 RESULT 7
 AAR58903
 ID AAR58903 standard; Protein; 43 AA.
 XX
 AC AAR58903;
 XX
 DT 17-APR-1995 (first entry)
 XX
 DE Drosophila-13 cadherin-related molecule.
 XX
 KW Cadherin; cell adhesion molecule.
 XX
 OS Drosophila.
 XX
 PN WO9414960-A.
 XX
 PD 07-JUL-1994.
 XX
 PF 23-DEC-1993; 93WO-US12588.
 XX
 PR 29-DEC-1992; 92US-0998003.
 XX
 PA (DOHE-) DOHENY EYE INST.
 XX
 PI Suzuki S;
 XX
 DR WPI; 1994-293849/36.
 DR N-PSDB; AAR58903.
 XX
 PT Polynucleotide sequences encoding new proto:cadherins - useful
 PT for modulating natural binding and regulating activities.
 XX
 PS Example; Page 64; 114pp; English.
 XX
 CC Two regions of conserved AA sequence, one from the middle of the
 CC third cadherin extracellular subdomain (EC-3) and the other from the
 CC C-terminus of the fourth extracellular subdomain (EC-4) were
 CC identified. The corresp. degenerate oligos (AAQ68949, AAQ68950) were
 CC designed for use as PCR primers. PCR was carried out on a rat brain
 CC cDNA prep. Two major bands of about 450 bps and 130 bps were found.
 CC The 450 bp band corresponded to the expected length between the two
 CC primer sites, but the 130 bp band could not be predicted from any
 CC of the previously identified cadherin sequences. The 450 bp and 130
 CC bp bands were extracted and sequenced. Nineteen novel partial cDNA
 CC clones (including sequences corresp. to the PCR primers) are given
 CC in AAQ68951-Q68969 and AAR58860-R58878. Various cDNA fragments
 CC structurally similar to the rat cDNAs were isolated from human,
 CC mouse and Xenopus brain cDNA preps. and from Drosophila and
 CC C. elegans whole body cDNA preps. by PCR using the above primers.
 CC The DNA and deduced AA sequences of the resulting PCR fragments
 CC (including sequences corresp. to the PCR primers) are given in
 CC AAQ68971, AAQ68972-Q68994 and AAR58882-R58905 and AAR49143. Comparison
 CC of the deduced AA sequences indicates a similarity, in particular,
 CC there are three sets of clones that appear to be cross-species
 CC homologues: RAT-218, MOUSE-322 and HUMAN-43; RAT-314, MOUSE-321 and
 CC HUMAN-11; and MOUSE-326 and HUMAN-42.
 XX
 SQ Sequence 43 AA;
 Query Match 96.0%; Score 24; DB 15; Length 43;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 |
 |
 |
 DB 37 DANDN 41

RESULT 8

AAR87102
ID AAR87102 standard; Peptide; 43 AA.

XX AAR87102;
XX 28-AUG-1996 (first entry)
XX Protocadherin clone RAT-123.
XX Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion;
KW catenin; therapy; clone.
XX Rattus rattus.
XX WO9600289-A1.
XX 04-JAN-1996.
XX 26-JUN-1995; 95WO-US08071.
XX 27-JUN-1994; 94US-0268161.
XX (DOHE-) DOHENY EYE INST.
XX Suzuki S;
XX WPI; 1996-068873/07.
XX N-PSDB; AAT03577.
XX Polynucleotide(s) encoding human protocadherins pc3 and pc4 and rat
PT pc5 - involved in cell-cell adhesion and regulation activities
XX Example 1; Page 36; 146pp; English.

XX AAR87102-R87120 represent partial fragments of the rat protocadherin
CC sequence. The cDNAs encoding these sequences were isolated after
CC screening a rat brain cDNA preparation with the primers shown in
CC AAT03575 and AAT03576. The primers were constructed from portions of the
CC amino acid sequences of the third and fourth extracellular domains of
CC published cadherin sequences. The full length cDNA sequence encoding
CC rat protocadherin-pc5 is represented in AAT03574. The cytoplasmic domain
CC of cadherin interacts with the cytoskeleton through catenins and other
CC cytoskeleton associated proteins. The cytoplasmic domain is not present
CC in all cadherins, but in those which possess it, it is essential for the
CC cadherins adhesive function. The cadherins which do not possess a
CC cytoplasmic domain appear to function via a different method from those
CC with a cytoplasmic domain. These protein sequences are involved in
CC cell-cell adhesion. These sequences may have regulatory functions in
CC the cell, as well as the cell-cell adhesive properties. Antibodies
CC produced against these sequences are useful for modulating the binding
CC activity of these protocadherins, and can be used therapeutically.

SQ Sequence 43 AA;

Query Match 96.0%; Score 24; DB 17; Length 43;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
| | | | |
Db 37 DANDN 41

RESULT 9
AAR87143
ID AAR87143 standard; Peptide; 43 AA.

XX AAR87143;
XX 29-AUG-1996 (first entry)
XX Protocadherin clone DROSOPHILA-13.
XX

KW Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion; mouse;
XX catenin; therapy; clone; frog; fruit fly.
XX Drosophila melanogaster.
XX WO9600289-A1.
XX 04-JAN-1996.
XX 26-JUN-1995; 95WO-US08071.
XX 27-JUN-1994; 94US-0268161.
XX (DOHE-) DOHENY EYE INST.
XX Suzuki S;
XX WPI; 1996-068873/07.
XX N-PSDB; AAT03618.
XX Polynucleotide(s) encoding human protocadherins pc3 and pc4 and rat
PT pc5 - involved in cell-cell adhesion and regulation activities
XX Example 2; Page 68; 146pp; English.

XX AAR87142-R87144 represent partial fragments of the drosophila
CC protocadherin sequence. The cDNAs encoding these sequences were isolated
CC after screening a drosophila whole body cDNA preparation with the primers
CC shown in AAT03575 and AAT03576. The primers were constructed from
CC portions of the amino acid sequences of the third and fourth
CC extracellular domains of published cadherin sequences. The cytoplasmic
CC domain of cadherin interacts with the cytoskeleton through catenins and
CC other cytoskeleton associated proteins. The cytoplasmic domain is not
CC present in all cadherins, but in those which possess it, it is essential
CC for the cadherins adhesive function. The cadherins which do not possess
CC a cytoplasmic domain appear to function via a different method from
CC those with a cytoplasmic domain. These protein sequences are involved
CC in cell-cell adhesion. These sequences may have regulatory functions in
CC the cell, as well as the cell-cell adhesive properties. Antibodies
CC produced against these sequences are useful for modulating the binding
CC activity of these protocadherins, and can be used therapeutically.

SQ Sequence 43 AA;

Query Match 96.0%; Score 24; DB 17; Length 43;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
| | | | |
Db 37 DANDN 41

RESULT 10
AAR46073
ID AAR46073 standard; Protein; 45 AA.

XX AAR46073;
XX 19-OCT-1994 (first entry)
XX Cadonectin like protein.
XX Human cDNA; library; enzyme; protein.
XX Homo sapiens.
XX WO9403599-A.
XX 17-FEB-1994.
XX 04-AUG-1993; 93WO-JP01095.
XX

PR 04-AUG-1992; 92JP-0208077.
 PR 13-NOV-1992; 92JP-0327619.
 PR 26-FEB-1993; 93JP-0061431.
 XX
 PA (SAGA) SAGAMI CHEM RES CENTRE.
 XX
 PI Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;
 XX
 DR WPI; 1994-065688/08.
 DR N-PSDB; AAQ57412.
 XX
 PT CDNA of human origin and proteins coded by it - which may be
 PT expressed by in vivo or in vitro translation using sense RNA or
 PT antisense DNA corresponding to the cDNA.
 XX
 PS Claim 1; Page 27; 167pp; Japanese.
 XX
 CC mRNA expressed in human fibrosarcoma cell line HT-1080 was
 CC isolated and used to construct a cDNA library using vector
 CC pKAL. Clone HP00012 encoding cadonectin-like protein
 CC was isolated.
 XX
 SQ Sequence 45 AA;
 Query Match 96.0%; Score 24; DB 15; Length 45;
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 5 DANDN 9
 RESULT 11
 ABP32657
 ID ABP32657 standard; Protein; 60 AA.
 AC
 AC ABP32657;
 XX
 DT 09-JUL-2002 (first entry)
 XX
 DE Human cadherin-like ORF1630 protein, SEQ ID NO:3260.
 XX
 KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipruritic; antidiabetic; cytostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US17076.
 XX
 PR 24-MAY-2000; 2000US-206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shinkets RA;
 XX
 XX WPI; 2002-106200/14.
 DR N-PSDB; ABN76683.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation
 XX
 PS Claim 10; Page 1074; 2508pp; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN75987 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, tumour inhibition activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins, and
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

SQ Sequence 60 AA;
 Query Match 96.0%; Score 24; DB 23; Length 60;
 Best Local Similarity 80.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 47 DANDN 51
 RESULT 12
 AAY70433
 ID AAY70433 standard; Protein; 81 AA.
 XX
 AC AAY70433;
 XX
 DT 21-JUN-2000 (first entry)
 XX
 DE Haematobia irritans thrombostatin protein.
 XX
 KW Thrombostatin; horn fly; thrombin; antithrombin; thrombolytic;
 KW coagulation; veterinary vaccine; treatment; haematophagy; prophylaxis;
 KW thromboembolism; thrombosis; acute shock therapy; coagulopathies;
 KW haemodialysis; haemoseparation; extracorporeal blood circulation;
 KW anticoagulant.
 XX
 OS Haematobia irritans.
 XX
 PN WO200011172-A1.

XX PD 02-MAR-2000.
 XX PF 19-AUG-1999; 99WO-US18888.
 XX PR 20-AUG-1998; 98US-0097227.
 XX PA (UYAU-) UNIV AUBURN.
 XX PI Cupp EW, Cupp MS;
 XX DR WPI; 2000-246563/21.
 XX N-PSDB; AA251574.
 XX PT Proteins with antithrombin activity useful as veterinary vaccines for
 PT prevention and treatment of hemostopathy in cattle and thrombosis in
 PT mammals are isolated from the salivary glands of horn fly -
 XX PS Claim 4b; Page 52; 58pp; English.
 XX CC The present amino acid sequence is the protein having antithrombin
 CC activity, designated as thrombostatin. It is isolated from the salivary
 CC glands of Haematobia irritans (horn fly), that contains an inhibitor of
 CC thrombin. The protein prevents coagulation of blood by inhibiting the
 CC activity of thrombin (factor II). It comprises 21% of aspartic acid and
 CC glutamic acid residues. This sequence has thrombolytic activity.
 CC Thrombostatin sequences are useful as veterinary vaccines for treating
 CC haemostopathy (blood-feeding) in cattle and thrombosis in mammals. It is
 CC useful for therapy and prophylaxis of thrombosis and thromboembolisms in
 CC humans, including prophylaxis of post-operative thrombosis, acute shock
 CC therapy, therapy for consumption of coagulopathies, in haemodialyses,
 CC haemoseparations and extracorporeal blood circulation. It can also be
 CC used as an anticoagulant in blood.
 XX SQ Sequence 81 AA;
 Query Match 96.0%; Score 24; DB 21; Length 81;
 Best Local Similarity 80.0%; Pred. No. 4.3e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 30 DSNDN 34
 RESULT 13
 ABB42873
 ID ABB42873 standard; Peptide; 86 AA.
 AC ABB42873;
 XX DT 04-FEB-2002 (first entry)
 XX DE Peptide #10379 encoded by human foetal liver single exon probe.
 XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX OS Homo sapiens.
 XX PN WO200157277-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00669.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human foetal liver -
 XX PS Claim 27; SEQ ID NO 35508; 639pp + sequence listing; English.
 XX CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 86 AA;
 Query Match 96.0%; Score 24; DB 22; Length 86;
 Best Local Similarity 80.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 29 DTNDN 33
 RESULT 14
 ABB26145
 ID ABB26145 standard; Protein; 86 AA.
 XX AC ABB26145;
 XX DT 23-JAN-2002 (first entry)
 XX DE Protein #8144 encoded by probe for measuring heart cell gene expression.
 XX KW Human; gene expression; heart; microarray; vascular system;
 XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
 XX KW congenital heart disease.
 XX OS Homo sapiens.
 XX PN WO200157274-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00666.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts -
 XX PS Claim 15; SEQ ID No 27915; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 86 AA;

Query Match 96.0%; Score 24; DB 22; Length 86;
 Best Local Similarity 80.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DXNDN 5
 Db 29 DTNDN 33

RESULT 15

AAW63768
 ID AAW63768 standard; Protein; 86 AA.

XX AC AAW63768;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35873.
 XX KW Human; brain expressed exon; gene expression analysis; probe;
 XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 XX KW epilepsy; cancer.
 XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT brains -
 XX PS Example 4; SEQ ID NO: 35873; 650pp + Sequence Listing; English.
 XX CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX

SQ Sequence 86 AA;

Query Match 96.0%; Score 24; DB 22; Length 86;
 Best Local Similarity 80.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DXNDN 5
 Db 29 DTNDN 33

RESULT 16

AAW76581
 ID AAW76581 standard; Protein; 86 AA.

XX AC AAW76581;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36887.
 XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
 XX KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human bone marrow -
 XX PS Example 4; SEQ ID NO: 36887; 658pp + Sequence Listing; English.
 XX CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.

XX SQ Sequence 86 AA;

Query Match 96.0%; Score 24; DB 22; Length 86;
 Best Local Similarity 80.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DXNDN 5
 Db 29 DTNDN 33

RESULT 17

AAW36690
 ID AAW36690 standard; Protein; 86 AA.

XX

AC AAM36690;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #10727 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
XX Homo sapiens.
OS
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PP 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
PT
XX Claim 27; SEQ ID No 36959; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP;
CC see AI131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 86 AA;
Query Match 96.0%; Score 24; DB 22; Length 86;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DXNDN 5
Db 29 DTNDN 33
RESULT 18
AAU19683
ID AAU19683 standard; Protein; 91 AA.
XX
AC AAU19683;
XX
XX 06-DEC-2001 (first entry)
DT
XX Human novel extracellular matrix protein, Seq ID No 333.
DE
XX Human; secreted extracellular matrix protein; immunomodulatory;
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW

KW wound healing; immunogen; gene therapy; antisense; food additive.
XX Homo sapiens.
OS
XX WO20015368-A1.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01348.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228524.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX

(HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465572/50.

XX N-PSDB; AAS31254.

XX Nucleic acid molecules encoding human secreted extracellular matrix proteins, used in preventing, treating or ameliorating a disorder, e.g.

XX Alzheimer's and Parkinson's diseases and cancers -

XX Claim 11; SEQ ID NO 333; 577pp; English.

XX The invention relates to isolated nucleic acid molecules encoding novel human secreted extracellular matrix proteins (SPs). The polynucleotides and proteins are used to prevent, treat a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For example, disorders associated with decreased expression of SPs. The SP polynucleotide or a vector expressing them may be administered to treat diseases by gene therapy. Antisense molecules may be administered to down regulate expression of SPs by binding with the cells own genes and preventing their expression. The polynucleotides may also be used as DNA probes in diagnostic assays. The SPs may also be used as antigens to produce antibodies and to identify modulators (agonists and antagonists) of the SPs. The anti-(SP) antibodies and antagonists may also be used to down regulate expression and activity of SP and as diagnostic agents for detecting the presence of SPs in samples. The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). Other uses include wound healing, maintenance of organs before transplantation.

Query Match 96.0%; Score 24; DB 22; Length 91;

Best Local Similarity 80.0%; Pred. No. 4.9e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5

Db 5 DANDN 9

RESULT 19

ABP47903

ID ABP47903 standard; Protein; 91 AA.

XX AC ABP47903;

XX DT 23-AUG-2002 (first entry)

XX DE Human polypeptide SEQ ID NO 333.

XX Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery; antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective; antiinflammatory; KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal; KW antiparasitic; cardiant; immune disorder; cardiovascular disorder; KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX OS Homo sapiens.

XX US2002042386-A1.

XX PD 11-APR-2002.

XX PF 17-JAN-2001; 2001US-0764870.

PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234937P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.

(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2002-470713/50.
N-PSDB; ABQ66578.

New nucleic acid encoding human proteins, useful for diagnosis, treatment and prevention of e.g. osteoporosis, also related polypeptides and antibodies

Claim 11; SEQ ID NO 333; 235pp + Sequence Listing; English.

The invention relates to novel genes (ABQ66521-ABQ66785) and proteins (ABP47846-ABP48110) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=99909764870.
XX
SQ Sequence 91 AA;

Query Match 96.0%; Score 24; DB 23; Length 91;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
| | | |
Db 5 DANDN 9

RESULT 20

AAV64614

ID AAY64614 standard; Peptide; 107 AA.

XX AAY64614;

XX 02-MAR-2000 (first entry)

XX Nonclassical cadherin extracellular domain SEQ ID NO:42.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW Inhibition; cadherin extracellular domain; cell adhesion recognition;
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW cadherin related neuronal receptor; Li-cadherin; protocadherin;
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW neurological disease.

XX Mammalia.

XX WO9957149-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00363.

XX 05-MAY-1998; 98US-0073040.

XX 06-NOV-1998; 98US-0187859.

XX 20-JAN-1999; 98US-0234395.

XX 08-MAR-1999; 99US-0264516.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-038791/03.

XX New cadherin modulating agents, used for modulating nonclassical
PT cadherin-mediated functions for treating e.g. cancers, obesity,
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
PT disease

XX Disclosure; Fig 2; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)
CC comprising peptides which comprise a nonclassical cadherin cell adhesion
CC recognition (CAR) sequence. The MA can be used for modulating
CC nonclassical cadherin-mediated functions. They can be used for e.g.
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a

CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
 CC in a mammal. They can also be used for treating e.g. psoriasis,
 CC arthritis, age-related macular degeneration, multiple sclerosis and
 CC diabetes. The products can also be used for detection and diagnosis and
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent
 CC sequences used in the exemplification of the present invention.

XX SQ Sequence 107 AA;

Query Match 96.0%; Score 24; DB 21; Length 107;
 Best Local Similarity 80.0%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DXNDN 5
 | | | |
 DB 99 DTNDN 103

RESULT 21

AAW04176
 ID AAW04176 standard; Protein; 120 AA.

AC AAW04176;

DT 19-MAY-1997 (first entry)

DE Variant variable heavy chain of Fas ligand antibody NOK-1.

KW Variable region; heavy chain; human; Fas ligand; monoclonal;
 KW antibody; NOK-1; hybridoma; inhibition; apoptosis; assay;
 KW diagnosis; disease; hepatitis; infectious mononucleosis;
 KW systemic lupus erythematosus; variant.

XX OS Mus musculus.

XX PN WO9629350-A1.

XX PD 26-SEP-1996.

XX PF 21-MAR-1996; 96WO-JP00734.

XX PR 27-OCT-1995; 95JP-0303492.

XX PR 20-MAR-1995; 95JP-0087420.

XX (SUME) SUMITOMO ELECTRIC IND CO.

XX PI Kayagaki N, Nakata M, Okumura K, Yagita H;

XX DR WPI; 1996-443140/44.

XX DR N-PSDB; AAT39549.

XX Monoclonal antibody specifically recognising the Fas ligand - useful
 PT for the detection of Fas ligands either on cell surface or in
 PT solution

XX PS Claim 19; Page 79; 133pp; Japanese.

XX The present sequence is a variant heavy chain variable region
 CC of the anti-human Fas ligand monoclonal antibody (Mab) NOK-1.
 CC It is produced by the hybridoma NOK-1 (FERM BP-5044), which was
 CC prepared by immunising mice with transformed human Fas ligand

CC expressing COS cells, and fusing spleen cells isolated from the
 CC mice with myeloma P3x63Ag8.653 (ATCC CRL-1580) cells. The Mab
 CC recognises the human Fas ligand on the cell surface or in solution,
 CC and can be used to inhibit the apoptosis inducing cell surface Fas
 CC ligand/Fas reaction. The Mab can also be used for a Fas ligand
 CC assay in biological samples (e.g. human blood), especially for
 CC disease diagnosis, e.g. hepatitis, infectious mononucleosis and
 CC systemic lupus erythematosus.

XX SQ Sequence 120 AA;

Query Match 96.0%; Score 24; DB 17; Length 120;
 Best Local Similarity 80.0%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DXNDN 5
 | | | |
 DB 56 DTNDN 60

RESULT 22

AAW16619
 ID AAW16619 standard; Protein; 120 AA.

AC AAW16619;

DT 14-JAN-1998 (first entry)

DE Anti-human FasL antibody (NOK1) heavy chain variable region.

KW Heavy chain; variable region; mouse; murine; human; Fas ligand;
 KW FasL; monoclonal antibody; Mab; hybridoma; treatment; hepatitis;
 KW hepatitis B virus; HBV; hepatitis C virus; HCV; apoptosis;
 KW liver cell; glutamate oxaloacetate; pyruvate transaminase.

XX OS Mus sp.

XX PN WO9715326-A1.

XX PD 01-MAY-1997.

XX PF 24-OCT-1996; 96WO-JP03089.

XX PR 27-OCT-1995; 95JP-0303491.

XX (SUME) SUMITOMO ELECTRIC IND CO.

XX PI Kayagaki N, Nakata M, Okumura K, Seino K, Yagita H;

XX DR WPI; 1997-258767/23.

XX DR N-PSDB; AAT66709.

XX Anti-human Fas Ligand antibody to treat hepatitis - controls
 PT apoptosis in liver cells and improves liver function

XX PS Claim 6; Page 29; 51pp; Japanese.

XX The present sequence is the heavy chain variable region of the
 CC murine anti-human Fas ligand (FasL) monoclonal antibody (Mab) NOK1,
 CC which is expressed by the hybridoma NOK1 (FERM BP-5044). The Mab
 CC can be used in the preparation of a composition for the effective
 CC oral or parenteral treatment of hepatitis, including hepatitis
 CC caused by hepatitis B or C virus. The composition controls apoptosis
 CC in liver cells caused by the binding of FasL to Fas expressing liver
 CC cells, and improves liver function by improving blood glutamate
 CC oxaloacetate and pyruvate transaminase levels. The composition is
 CC given in a dosage of 0.0001-1000, preferably 0.01-600 mg/day.
 CC Spleen cells from mice immunised with FasL expressing COS cells
 CC were fused with mouse myeloma cells to produce hybridomas. The
 CC hybridomas were screened for anti-FasL activity, and the active
 CC clones NOK1-5 isolated.

XX SQ Sequence 120 AA;

Query Match 96.0%; Score 24; DB 18; Length 120;
 Best Local Similarity 80.0%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 56 DTNDN 60

RESULT 23

AAW00833
 ID AAW00833 standard; Protein; 121 AA.

XX AC AAW00833;

XX DT 20-MAY-1997 (first. entry)

XX DE Variable heavy chain of anti-human Fas ligand antibody NOK-1.

XX KW Variable region; heavy chain; human; Fas ligand; monoclonal;
 KW antibody; NOK-1; hybridoma; inhibition; apoptosis; assay;
 KW diagnosis; disease; hepatitis; infectious mononucleosis;
 KW systemic lupus erythematosus.

XX OS Mus musculus.

XX PH Key Location/Qualifiers

XX FT Misc-difference 49 /note= "corresponding codon absent"

XX PN WO9629350-A1.

XX PD 26-SEP-1996.

XX PF 21-MAR-1996; 96WO-JP00734.

XX PR 27-OCT-1995; 95JP-0303492.

XX PR 20-MAR-1995; 95JP-0087420.

XX PA (SUME) SUMITOMO ELECTRIC IND CO.

XX PI Kayagaki N, Nakata M, Okumura K, Yagita H;

XX DR WPI: 1996-443140/44.

XX DR N-PSDB: RAT39559.

XX PT Monoclonal antibody specifically recognising the Fas ligand - useful
 PT for the detection of Fas ligands either on cell surface or in
 PT solution

XX PS Claim 40; Page 92; 133pp; Japanese.

XX CC The present sequence is the heavy chain variable region of the
 CC anti-human Fas ligand monoclonal antibody (MAB) NOK-1. NOK-1 is
 CC produced by the hybridoma NOK-1 (FERM BP-5044), which was prepared
 CC by immunising mice with transformed human Fas ligand expressing COS
 CC cells, and fusing spleen cells isolated from the mice with myeloma
 CC P3x63Ag8.653 (ATCC CRL-1580) cells. The MAB recognises the human
 CC Fas ligand on the cell surface or in solution, and can be used to
 CC inhibit the apoptosis inducing cell surface Fas ligand/Fas
 CC reaction. The MAB can also be used for a Fas ligand assay in
 CC biological samples (e.g. human blood), especially for disease
 CC diagnosis, e.g. hepatitis, infectious mononucleosis and systemic
 CC lupus erythematosus.

XX SQ Sequence 121 AA;

Query Match 96.0%; Score 24; DB 17; Length 121;
 Best Local Similarity 80.0%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5

Db 57 DTNDN 61

RESULT 24

ABP08216
 ID ABP08216 standard; Protein; 131 AA.

XX AC ABP08216;

XX DT 24-JUN-2002 (first entry)

XX DE Human ORFX protein sequence SEQ ID NO:16414.

XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX OS Homo sapiens.

XX PN WO200192523-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US10836.

XX PR 30-MAY-2000; 2000US-206132P.

XX PR 29-AUG-2000; 2000US-228716P.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach MD;

XX DR WPI: 2002-106308/14.

XX DR N-PSDB: ABN23968.

XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 PS Disclosure; SEQ ID 16414; 1037pp; English.

XX CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.

XX CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 131 AA;

Query Match 96.0%; Score 24; DB 23; Length 131;
 Best Local Similarity 80.0%; Pred. No. 7.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DXNDN 5
 Db 78 DSDN 82

RESULT 25

AAO05889
 ID AAO05889 standard; Protein; 134 AA.

XX AC AAO05889;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 19781.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US04927.

XX PR 28-FEB-2000; 2000US-0515126.

XX PR 18-MAY-2000; 2000US-0577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AAI85820.

XX PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders

XX PS Claim 20; SEQ ID NO 19781; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

XX CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 134 AA;

Query Match 96.0%; Score 24; DB 22; Length 134;
 Best Local Similarity 80.0%; Pred. No. 7.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DXNDN 5
 Db 78 DSDN 82

Db 22 DANDN 26

Search completed: June 20, 2003, 20:51:48
 Job time : 36.9318 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 10.5682 Seconds
(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-8
Perfect score: 25
Sequence: 1 DXNDN 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	96.0	15	1	US-08-167-336A-7
2	24	96.0	15	1	US-08-416-962-7
3	24	96.0	15	1	US-08-589-011-7
4	24	96.0	15	2	US-08-948-762-7
5	24	96.0	20	4	US-08-899-279-13
6	24	96.0	20	4	US-08-899-279-13
7	24	96.0	40	1	US-07-998-003A-55
8	24	96.0	40	1	US-08-453-274B-55
9	24	96.0	40	1	US-08-453-695A-55
10	24	96.0	40	1	US-08-268-161A-55
11	24	96.0	40	2	US-08-453-702A-55
12	24	96.0	40	4	US-09-099-639-55
13	24	96.0	40	5	PCT-US93-12588-55
14	24	96.0	40	5	PCT-US95-08071-55
15	24	96.0	43	1	US-07-998-003A-89
16	24	96.0	43	1	US-08-453-274B-89
17	24	96.0	43	1	US-08-453-695A-89
18	24	96.0	43	1	US-08-453-702A-89
19	24	96.0	43	1	US-08-453-695A-89
20	24	96.0	43	1	US-08-453-695A-89
21	24	96.0	43	1	US-08-268-161A-89
22	24	96.0	43	2	US-08-453-702A-89
23	24	96.0	43	2	US-08-453-702A-89
24	24	96.0	43	2	US-08-453-702A-89
25	24	96.0	43	4	US-09-099-639-89
26	24	96.0	43	4	US-09-099-639-89
27	24	96.0	43	5	PCT-US93-12588-4

28	24	96.0	43	5	PCT-US93-12588-89	Sequence 89, Appl
29	24	96.0	43	5	PCT-US95-08071-4	Sequence 4, Appl
30	24	96.0	43	5	PCT-US95-08071-89	Sequence 89, Appl
31	24	96.0	81	4	US-09-376-113-2	Sequence 2, Appl
32	24	96.0	92	4	US-09-178-176B-10	Sequence 10, Appl
33	24	96.0	92	4	US-09-457-864-10	Sequence 10, Appl
34	24	96.0	93	1	US-08-326-117B-5	Sequence 5, Appl
35	24	96.0	93	3	US-08-982-129-5	Sequence 5, Appl
36	24	96.0	107	4	US-09-187-859-42	Sequence 42, Appl
37	24	96.0	120	3	US-05-065-059-1	Sequence 1, Appl
38	24	96.0	146	1	US-07-998-003A-38	Sequence 38, Appl
39	24	96.0	146	1	US-08-453-274B-38	Sequence 38, Appl
40	24	96.0	146	1	US-08-453-695A-38	Sequence 38, Appl
41	24	96.0	146	1	US-08-268-161A-38	Sequence 38, Appl
42	24	96.0	146	2	US-08-453-702A-38	Sequence 38, Appl
43	24	96.0	146	4	US-09-099-639-38	Sequence 38, Appl
44	24	96.0	146	5	PCT-US93-12588-38	Sequence 38, Appl
45	24	96.0	146	5	PCT-US95-08071-38	Sequence 38, Appl
46	24	96.0	168	4	US-09-376-113-5	Sequence 5, Appl
47	24	96.0	175	4	US-09-376-113-7	Sequence 7, Appl
48	24	96.0	237	1	US-08-750-532-18	Sequence 18, Appl
49	24	96.0	332	4	US-09-134-001C-4646	Sequence 4646, Ap
50	24	96.0	367	2	US-08-515-251A-4	Sequence 4, Appl
51	24	96.0	444	4	US-09-134-001C-4346	Sequence 4346, Ap
52	24	96.0	493	6	5486473-6	Patent No. 5486473
53	24	96.0	533	4	US-09-221-275-4	Sequence 4, Appl
54	24	96.0	582	4	US-09-242-632A-2	Sequence 2, Appl
55	24	96.0	616	1	US-08-453-695A-115	Sequence 115, App
56	24	96.0	616	1	US-08-268-161A-115	Sequence 115, App
57	24	96.0	616	2	US-08-453-702A-115	Sequence 115, App
58	24	96.0	616	4	US-09-099-639-115	Sequence 115, App
59	24	96.0	616	5	PCT-US95-08071-115	Sequence 115, App
60	24	96.0	659	4	US-08-894-818B-5	Sequence 5, Appl
61	24	96.0	659	4	US-08-894-818B-5	Sequence 5, Appl
62	24	96.0	659	4	US-09-445-472-12	Sequence 12, Appl
63	24	96.0	682	1	US-07-998-003A-107	Sequence 107, App
64	24	96.0	682	1	US-08-453-274B-107	Sequence 107, App
65	24	96.0	682	1	US-08-453-695A-107	Sequence 107, App
66	24	96.0	682	1	US-08-268-161A-107	Sequence 107, App
67	24	96.0	682	2	US-08-453-702A-107	Sequence 107, App
68	24	96.0	682	4	US-09-099-639-107	Sequence 107, App
69	24	96.0	682	5	PCT-US93-12588-107	Sequence 107, App
70	24	96.0	682	5	PCT-US95-08071-107	Sequence 107, App
71	24	96.0	694	2	US-08-701-240-2	Sequence 2, Appl
72	24	96.0	694	4	US-09-138-236-2	Sequence 2, Appl
73	24	96.0	747	3	US-09-035-648-18	Sequence 18, Appl
74	24	96.0	747	4	US-09-001-951-18	Sequence 18, Appl
75	24	96.0	747	4	US-08-818-829-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-167-336A-7
; Sequence 7, Application US/08167336A
; Patent No. 5531990
; GENERAL INFORMATION:
; APPLICANT: THANAVALA, YASMIN
; APPLICANT: THAKUR, ARVIND
; APPLICANT: ROITT, IVAN
; APPLICANT: PRIDE, MICHAEL
; TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY
; TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS
; TITLE OF INVENTION: B SURFACE ANTIGEN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DUNN & ASSOCIATES, P.C.
; STREET: P.O. BOX 96
; CITY: NEWFANE
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 14108

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
COMPUTER: VICTOR 300 SX/25
OPERATING SYSTEM: MS-DOS VERSION 5.0
SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,336A
FILING DATE: 15-DEC-1993
CLASSIFICATION: 530
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: DUNN, MICHAEL L.
REGISTRATION NUMBER: 25,330
REFERENCE/DOCKET NUMBER: RPP:138 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 433-1661
TELEFAX: (716) 433-1665
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
MOLECULE TYPE: PEPTIDE
HYPOTHETICAL: NO
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-167-336A-7

Query Match 96.0%; Score 24; DB 1; Length 15;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
Db 8 DTNDN 12

RESULT 2
US-08-416-962-7
Sequence 7, Application US/08416962
Patent No. 5668253
GENERAL INFORMATION:
APPLICANT: THANAVALA, YASMIN
APPLICANT: THAKUR, ARVIND
APPLICANT: ROITT, IVAN
APPLICANT: PRIDE, MICHAEL
TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY
TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS
TITLE OF INVENTION: B SURFACE ANTIGEN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: DUNN & ASSOCIATES, P.C.
STREET: P.O. BOX 96
CITY: NEWFANE
STATE: NEW YORK
COUNTRY: USA
ZIP: 14108
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
COMPUTER: VICTOR 300 SX/25
OPERATING SYSTEM: MS-DOS VERSION 5.0
SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,962
FILING DATE: 05-APR-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,336
FILING DATE: 15-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: DUNN, MICHAEL L.
REGISTRATION NUMBER: 25,330
REFERENCE/DOCKET NUMBER: RPP:138 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 433-1661
TELEFAX: (716) 433-1665
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
MOLECULE TYPE: PEPTIDE
HYPOTHETICAL: NO
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:

US-08-416-962-7

Query Match 96.0%; Score 24; DB 1; Length 15;

Best Local Similarity 80.0%; Pred. No. 26;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5

Db 8 DTNDN 12

RESULT 3

US-08-589-011-7

Sequence 7, Application US/08589011

Patent No. 5744135

GENERAL INFORMATION:

APPLICANT: THANAVALA, YASMIN

APPLICANT: THAKUR, ARVIND

APPLICANT: ROITT, IVAN

APPLICANT: PRIDE, MICHAEL

TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY

TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS

TITLE OF INVENTION: B SURFACE ANTIGEN

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: DUNN & ASSOCIATES, P.C.

STREET: P.O. BOX 96

CITY: NEWFARE

STATE: NEW YORK

COUNTRY: USA

ZIP: 14108

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB

COMPUTER: VICTOR 300 SX/25

OPERATING SYSTEM: MS-DOS VERSION 5.0

SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/589,011

FILING DATE: 19-JAN-1996

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/167,336

FILING DATE: 15-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: DUNN, MICHAEL L.

REGISTRATION NUMBER: 25,330

REFERENCE/DOCKET NUMBER: RPP:138 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 433-1661

TELEFAX: (716) 433-1665

INFORMATION FOR SEQ ID NO: 7;

SEQUENCE CHARACTERISTICS:

LENGTH: 15

TYPE: AMINO ACID

STRANDEDNESS: UNKNOWN

TOPOLOGY: UNKNOWN

MOLECULE TYPE: PEPTIDE

HYPOTHETICAL: NO

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM:

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

LIBRARY:

CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:

US-08-589-011-7

Query Match

96.0%; Score 24; DB 1; Length 15;

Best Local Similarity 80.0%; Pred. No. 26;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5

Db 8 DTNDN 12

RESULT 4

US-08-948-762-7

Sequence 7, Application US/08948762

Patent No. 5856087

GENERAL INFORMATION:

APPLICANT: THANAVALA, YASMIN

APPLICANT: THAKUR, ARVIND

APPLICANT: ROITT, IVAN

APPLICANT: PRIDE, MICHAEL

TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY

TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS

TITLE OF INVENTION: B SURFACE ANTIGEN

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: DUNN & ASSOCIATES, P.C.

STREET: P.O. BOX 96

CITY: NEWFARE

STATE: NEW YORK

COUNTRY: USA

ZIP: 14108

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB

COMPUTER: VICTOR 300 SX/25

OPERATING SYSTEM: MS-DOS VERSION 5.0

SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/948,762

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/589,011
FILING DATE: 19-JAN-1996
APPLICATION NUMBER: 08/167,336
FILING DATE: 15-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: DUNN, MICHAEL L.
REGISTRATION NUMBER: 25,330
REFERENCE/DOCKET NUMBER: RPP:138 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 433-1661
TELEFAX: (716) 433-1665
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
MOLECULE TYPE: PEPTIDE
HYPOTHETICAL: NO
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-08-948-762-7
Query Match 96.0%; Score 24; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 8 DTNDN 12

RESULT 5
US-08-899-279-13
Sequence 13, Application US/08899279
Patent No. 6344199
GENERAL INFORMATION:
APPLICANT: Hofler, Heinz
ADDRESSEE: Becker, Karl-Friedrich

APPLICANT: Krenmer, Elisabeth
APPLICANT: Eulitz, Manfred
APPLICANT: Schuhmacher, Christoph
TITLE OF INVENTION: Mutations of E Cadherin as a Basis for
the Diagnosis and Therapy of Human Malignant Tumours
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,279
FILING DATE: 23-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 196 29 938.1
FILING DATE: 24-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Heines, M. Henry
REGISTRATION NUMBER: 28,219
REFERENCE/DOCKET NUMBER: 080296-00000000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /note= "amino acid sequence at
positions 360 through 379 of "normal" human E-cadherin pr
US-08-899-279-13

Query Match 96.0%; Score 24; DB 4; Length 20;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 8 DTNDN 12

RESULT 6
US-08-899-279-13
Sequence 13, Application US/08899279
Patent No. 6447776
GENERAL INFORMATION:
APPLICANT: Hofler, Heinz
ADDRESSEE: Becker, Karl-Friedrich
Krenmer, Elisabeth
Eulitz, Manfred
Schuhmacher, Christoph
TITLE OF INVENTION: Mutations of E Cadherin as a Basis for
the Diagnosis and Therapy of Human Malignant Tumours
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California

;; COUNTRY: USA
;; ZIP: 94111-3834
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: 23-Jul-1997
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: DE 196 29 938.1
;; FILING DATE: 24-JUL-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Heines M. Henry
;; REGISTRATION NUMBER: 28,219
;; REFERENCE/DOCKET NUMBER: 080296-0000000US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1..20
;; OTHER INFORMATION: /note= "amino acid sequence at
;; positions 360 through 379 of "normal" human E-cadherin protein"
;; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-08-999-279-13

Query Match 96.0%; Score 24; DB 4; Length 20;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 8 DTNDN 12

RESULT 7
US-07-998-003A-55
;; Sequence 55, Application US/07998003A
;; Patent No. 5643781
;; GENERAL INFORMATION:
;; APPLICANT: Suzuki, Shintaro
;; TITLE OF INVENTION: Protocadherin Materials and Methods
;; NUMBER OF SEQUENCES: 107
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
;; ADDRESSEE: Bicknell
;; STREET: 20 South Clark Street
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60603
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/998,003A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5643781and, Greta E.
;; REGISTRATION NUMBER: 35,302

;; REFERENCE/DOCKET NUMBER: 30903
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312/346-5750
;; TELEFAX: 312/984-9740
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 55:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 40 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-07-998-003A-55

Query Match 96.0%; Score 24; DB 1; Length 40;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 34 DANDN 38

RESULT 8
US-08-453-274B-55
;; Sequence 55, Application US/08453274B
;; Patent No. 5663300
;; GENERAL INFORMATION:
;; APPLICANT: Suzuki, Shintaro
;; TITLE OF INVENTION: Protocadherin Materials and Methods
;; NUMBER OF SEQUENCES: 107
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;; STREET: 6300 Sears Tower, 233 South Wacker Drive
;; CITY: Chicago.
;; STATE: Illinois
;; COUNTRY: United States of America
;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/453,274B
;; FILING DATE: 30-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5663300and, Greta E.
;; REGISTRATION NUMBER: 35,302
;; REFERENCE/DOCKET NUMBER: 32660
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312/474-6300
;; TELEFAX: 312/474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 55:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 40 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-453-274B-55

Query Match 96.0%; Score 24; DB 1; Length 40;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 34 DANDN 38

RESULT 9

US-08-453-695A-55
; Sequence 55, Application US/08453695A
; Patent No. 5708143
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,695A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5708143and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32658
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-453-695A-55

Query Match 96.0%; Score 24; DB 1; Length 40;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 1 DXNDN 5
Db 34 DANDN 38

RESULT 10
US-08-268-161A-55
; Sequence 55, Application US/08268161A
; Patent No. 5798224
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/268,161A

; FILING DATE: June 27, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Young J. Suh
; REGISTRATION NUMBER: P-41,337
; REFERENCE/DOCKET NUMBER: 27866/32149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-268-161A-55

Query Match 96.0%; Score 24; DB 1; Length 40;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 1 DXNDN 5
Db 34 DANDN 38

RESULT 11
US-08-453-702A-55
; Sequence 55, Application US/08453702A
; Patent No. 5891706
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,702A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5891706and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32657
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-453-702A-55

Query Match 96.0%; Score 24; DB 2; Length 40;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 1 DXNDN 5 ,
 Db 34 DANDN 38

RESULT 12

US-09-099-639-55
 ; Sequence 55, Application US/09099639
 ; Patent No. 6262237
 ; GENERAL INFORMATION:
 ; APPLICANT: Suzuki, Shintaro
 ; TITLE OF INVENTION: Protocadherin Materials and Methods
 ; NUMBER OF SEQUENCES: 115
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
 ; STREET: 233 South Wacker, 6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/099,639
 ; FILING DATE: 18 JUN 1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/263,161
 ; FILING DATE: 27 JUN 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Greta E. No. 6262237and
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 27866/34703
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 55:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 40 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-099-639-55

Query Match 96.0%; Score 24; DB 4; Length 40;
 Best Local Similarity 80.0%; Pred. No. 69;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 Db 34 DANDN 38

RESULT 13

PCT-US93-12588-55
 ; Sequence 55, Application PC/TUS9312588
 ; GENERAL INFORMATION:
 ; APPLICANT: Suzuki, Shintaro
 ; TITLE OF INVENTION: Protocadherin Materials and Methods
 ; NUMBER OF SEQUENCES: 107
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
 ; STREET: 6300 Sears Tower, 233 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA

ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/12588
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/998,003
 ; FILING DATE: 29 DEC 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Noland, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 31811
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 55:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 40 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US93-12588-55

Query Match 96.0%; Score 24; DB 5; Length 40;
 Best Local Similarity 80.0%; Pred. No. 69;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 Db 34 DANDN 38

RESULT 14

PCT-US95-08071-55
 ; Sequence 55, Application PC/TUS9508071
 ; GENERAL INFORMATION:
 ; APPLICANT: Suzuki, Shintaro
 ; TITLE OF INVENTION: Protocadherin Materials and Methods
 ; NUMBER OF SEQUENCES: 115
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gersbein, Murray, &
 ; STREET: 6300 Sears Tower, 233 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/08071
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/12588
 ; FILING DATE: 23 DEC 1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Noland, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 32149

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-08071-55

Query Match 96.0%; Score 24; DB 5; Length 40;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
Db 34 DANDN 38

RESULT 15
US-07-998-003A-4
; Sequence 4, Application US/07998003A
; Patent No. 5643781
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; STREET: 20 South Clark Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/998,003A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5643781and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 30903
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-998-003A-89

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
Db 37 DANDN 41

RESULT 16
US-07-998-003A-89
; Sequence 89, Application US/07998003A
; Patent No. 5643781
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; STREET: 20 South Clark Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/998,003A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5643781and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 30903
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-998-003A-89

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
Db 37 DANDN 41

RESULT 17
US-08-453-274B-4
; Sequence 4, Application US/08453274B
; Patent No. 5663300
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,274B

```

FILING DATE: 30-MAY-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5663300and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 32660
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 43 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-453-274B-4

Query Match 96.0%; Score 24; DB 1; Length 43;
 Best Local Similarity 80.0%; Pred. No. 74;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 37 DANDN 41

RESULT 18
 US-08-453-274B-89
 Sequence 89 Application US/08453274B
 Patent No. 5663300
 GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: Protocadherin Materials and Methods
 NUMBER OF SEQUENCES: 107
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 FILING DATE: 30-MAY-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5663300and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 32660
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 89:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 43 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-453-274B-89

Query Match 96.0%; Score 24; DB 1; Length 43;
 Best Local Similarity 80.0%; Pred. No. 74;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 37 DANDN 41

Db 37 DANDN 41

RESULT 19
 US-08-453-695A-4
 Sequence 4 Application US/08453695A
 Patent No. 5708143
 GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: Protocadherin Materials and Methods
 NUMBER OF SEQUENCES: 115
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
 ADDRESSEE: Borun
 STREET: 233 South Wacker, 6300 Sears Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 FILING DATE: US/08/453,695A
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5708143and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 32658
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 43 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-453-695A-4

Query Match 96.0%; Score 24; DB 1; Length 43;
 Best Local Similarity 80.0%; Pred. No. 74;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 37 DANDN 41

RESULT 20
 US-08-453-695A-89
 Sequence 89 Application US/08453695A
 Patent No. 5708143
 GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: Protocadherin Materials and Methods
 NUMBER OF SEQUENCES: 115
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
 ADDRESSEE: Borun
 STREET: 233 South Wacker, 6300 Sears Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,695A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 5708143and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32658
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-695A-89

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

RESULT 21
US-08-268-161A-4
Sequence 4, Application US/08268161A
Patent No. 5798224
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,161A
FILING DATE: June 27, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Young J. Suh
REGISTRATION NUMBER: P-41,337
REFERENCE/DOCKET NUMBER: 27866/32149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-268-161A-4

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

RESULT 22
US-08-268-161A-89
Sequence 89, Application US/08268161A
Patent No. 5798224
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,161A
FILING DATE: June 27, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Young J. Suh
REGISTRATION NUMBER: P-41,337
REFERENCE/DOCKET NUMBER: 27866/32149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-268-161A-89

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

RESULT 23
US-08-453-702A-4
Sequence 4, Application US/08453702A
Patent No. 5891706
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,161A
FILING DATE: June 27, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Young J. Suh
REGISTRATION NUMBER: P-41,337
REFERENCE/DOCKET NUMBER: 27866/32149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-268-161A-4

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

RESULT 22
US-08-268-161A-89
Sequence 89, Application US/08268161A
Patent No. 5798224
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,161A
FILING DATE: June 27, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Young J. Suh
REGISTRATION NUMBER: P-41,337
REFERENCE/DOCKET NUMBER: 27866/32149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-268-161A-89

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

RESULT 23
US-08-453-702A-4
Sequence 4, Application US/08453702A
Patent No. 5891706
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,161A
FILING DATE: June 27, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Young J. Suh
REGISTRATION NUMBER: P-41,337
REFERENCE/DOCKET NUMBER: 27866/32149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-268-161A-4

Query Match 96.0%; Score 24; DB 1; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60606
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA: US/08/453,702A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5891706and, Greta E.
;; REGISTRATION NUMBER: 35,302
;; REFERENCE/DOCKET NUMBER: 32657
;; TELEPHONE: 312/474-6300
;; TELEFAX: 312/474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-453-702A-4

Query Match 96.0%; Score 24; DB 2; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

RESULT 24
US-08-453-702A-89
; Sequence 89, Application US/08453702A
; Patent No. 5891706 ,
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5891706and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32657
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 25-3856
; INFORMATION FOR SEQ ID NO: 89:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-453-702A-89

Query Match 96.0%; Score 24; DB 2; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

RESULT 25
US-09-099-639-4
; Sequence 4, Application US/09099639
; Patent No. 6262237
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/099,639
; FILING DATE: 18 JUN 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/263,161
; FILING DATE: 27 JUN 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Greta E. No. 6262237and
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/34703
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-099-639-4

Query Match 96.0%; Score 24; DB 4; Length 43;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 37 DANDN 41

Search completed: June 20, 2003, 21:00:26
Job time : 12.5682 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:52:19 ; Search time 12.8409 Seconds
(without alignments)
42.13% Million cell updates/sec

Title: US-10-105-008-8

Perfect score: 25

Sequence: 1 DXNDN 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	96.0	20	9 US-10-047-403-13	Sequence 13, Appl
2	24	96.0	86	10 US-09-864-761-4143	Sequence 4143, A
3	24	96.0	91	9 US-10-125-540-333	Sequence 333, App
4	24	96.0	91	10 US-09-764-870-333	Sequence 333, App
5	24	96.0	92	9 US-10-187-088-10	Sequence 10, Appl
6	24	96.0	93	9 US-09-457-865-5	Sequence 5, Appli
7	24	96.0	107	9 US-10-006-869-42	Sequence 42, Appl
8	24	96.0	155	9 US-10-001-883-113	Sequence 113, App
9	24	96.0	165	9 US-10-125-540-483	Sequence 483, App
10	24	96.0	165	10 US-09-764-870-483	Sequence 483, App
11	24	96.0	177	9 US-09-791-279-148	Sequence 148, App
12	24	96.0	212	9 US-09-738-626-5604	Sequence 5604, Ap
13	24	96.0	339	9 US-10-125-540-458	Sequence 458, App
14	24	96.0	339	10 US-09-764-870-458	Sequence 458, App
15	24	96.0	427	9 US-10-125-540-482	Sequence 482, App
16	24	96.0	427	10 US-09-764-870-482	Sequence 482, App
17	24	96.0	453	9 US-10-125-540-296	Sequence 296, App
18	24	96.0	453	10 US-09-764-870-296	Sequence 296, App
19	24	96.0	526	10 US-09-999-162-12	Sequence 12, Appl

RESULT 1

US-10-047-403-13
; Sequence 13, Application US/10047403
; Publication No. US20030054005A1
; GENERAL INFORMATION:
; APPLICANT: Hoffer, Heinz
; Becker, Karl-Friedrich
; Kremmer, Elisabeth
; Eulitz, Manfred
; Schuhmacher, Christoph

TITLE OF INVENTION: Mutations of E Cadherin as a Basis for
the Diagnosis and Therapy of Human Malignant Tumours

ALIGNMENTS

20	24	96.0	586	9	US-09-908-193-28	Sequence 28, Appl
21	24	96.0	587	9	US-09-893-519A-72	Sequence 72, Appl
22	24	96.0	589	10	US-09-972-086-2	Sequence 2, Appli
23	24	96.0	602	9	US-09-908-193-4	Sequence 4, Appli
24	24	96.0	615	9	US-09-908-193-26	Sequence 26, Appl
25	24	96.0	615	9	US-09-908-193-27	Sequence 27, Appl
26	24	96.0	627	9	US-10-050-704-125	Sequence 125, App
27	24	96.0	659	12	US-10-090-624-12	Sequence 12, Appl
28	24	96.0	672	9	US-09-991-936-1862	Sequence 1862, Ap
29	24	96.0	672	9	US-09-991-936-1865	Sequence 1865, Ap
30	24	96.0	673	9	US-09-759-130B-6	Sequence 6, Appli
31	24	96.0	681	9	US-09-826-115-20	Sequence 20, Appl
32	24	96.0	681	9	US-09-826-115-24	Sequence 24, Appl
33	24	96.0	774	10	US-09-833-435A-8	Sequence 8, Appli
34	24	96.0	783	10	US-09-864-761-38209	Sequence 38209, A
35	24	96.0	786	9	US-10-002-344A-221	Sequence 221, App
36	24	96.0	791	9	US-10-125-540-337	Sequence 337, App
37	24	96.0	791	10	US-09-764-870-337	Sequence 337, App
38	24	96.0	794	10	US-09-864-761-37663	Sequence 37663, A
39	24	96.0	798	9	US-09-978-295A-405	Sequence 405, App
40	24	96.0	798	9	US-09-978-697-405	Sequence 405, App
41	24	96.0	798	9	US-09-978-192A-405	Sequence 405, App
42	24	96.0	798	9	US-09-999-832A-405	Sequence 405, App
43	24	96.0	798	9	US-09-978-189-405	Sequence 405, App
44	24	96.0	798	9	US-10-227-884-132	Sequence 132, App
45	24	96.0	798	9	US-10-230-163-132	Sequence 132, App
46	24	96.0	798	9	US-09-978-608A-405	Sequence 405, App
47	24	96.0	798	9	US-10-218-631-132	Sequence 132, App
48	24	96.0	798	9	US-10-230-338-132	Sequence 132, App
49	24	96.0	798	9	US-09-978-191A-405	Sequence 405, App
50	24	96.0	798	9	US-09-978-403A-405	Sequence 405, App
51	24	96.0	798	9	US-09-978-564A-405	Sequence 405, App
52	24	96.0	798	9	US-09-978-585A-405	Sequence 405, App
53	24	96.0	798	9	US-10-017-081A-405	Sequence 405, App
54	24	96.0	798	9	US-10-230-414-132	Sequence 132, App
55	24	96.0	798	9	US-09-978-824-405	Sequence 405, App
56	24	96.0	798	9	US-09-981-915A-405	Sequence 405, App
57	24	96.0	798	9	US-09-999-833A-405	Sequence 405, App
58	24	96.0	798	9	US-10-167-749-405	Sequence 405, App
59	24	96.0	798	9	US-09-918-585A-405	Sequence 405, App
60	24	96.0	798	9	US-09-978-423A-405	Sequence 405, App
61	24	96.0	798	9	US-10-013-921A-405	Sequence 405, App
62	24	96.0	798	9	US-10-216-159A-132	Sequence 132, App
63	24	96.0	798	9	US-09-978-193A-405	Sequence 405, App
64	24	96.0	798	9	US-10-013-929A-405	Sequence 405, App
65	24	96.0	798	9	US-10-016-177A-405	Sequence 405, App
66	24	96.0	798	9	US-10-218-849-132	Sequence 132, App
67	24	96.0	798	9	US-10-227-873-132	Sequence 132, App
68	24	96.0	798	9	US-10-227-883-132	Sequence 132, App
69	24	96.0	798	9	US-09-999-830A-405	Sequence 405, App
70	24	96.0	798	9	US-10-219-076-132	Sequence 132, App
71	24	96.0	798	9	US-10-230-434-132	Sequence 132, App
72	24	96.0	798	9	US-09-978-757A-405	Sequence 405, App
73	24	96.0	798	9	US-10-219-003-132	Sequence 132, App
74	24	96.0	798	9	US-10-219-075-132	Sequence 132, App
75	24	96.0	798	9	US-10-219-464-132	Sequence 132, App

NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/10/047,403
FILING DATE: 14-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/899,279
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Heines, M. Henry
REGISTRATION NUMBER: 28,219
REFERENCE/DOCKET NUMBER: 080296-0000000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /note="amino acid sequence at positions 360 through 379 of "normal" human E-cadherin protein"
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-047-403-13

Query Match 96.0%; Score 24; DB 9; Length 20;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 8 DTNDN 12

RESULT 2
US-09-864-761-41443
Sequence 41443, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 41443
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011859.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.76
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT HIT: Q94131, EVALUE 1.50e+00
OTHER INFORMATION: EST_HUMAN HIT: A1917113.1, EVALUE 3.00e-32
US-09-864-761-41443

Query Match 96.0%; Score 24; DB 10; Length 86;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 29 DTNDN 33

RESULT 3
US-10-125-540-333
Sequence 333, Application US/10125540
Publication No. US20030059875A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ14C1
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US/10/125,540
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 646
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 333
LENGTH: 91
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: misc feature
LOCATION: (43)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc_feature
LOCATION: (61)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (83)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-333

Query Match 96.0%; Score 24; DB 9; Length 91;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
|
|
|
|
Db 5 DANDN 9

RESULT 4
US-09-764-870-333
Sequence 333, Application US/09764870
Patent No. US20020042386A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PZ14
CURRENT APPLICATION NUMBER: US/09/764,870
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 646
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 333
LENGTH: 91
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (43)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (61)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (80)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (83)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-333

Query Match 96.0%; Score 24; DB 10; Length 91;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
|
|
|
|
Db 5 DANDN 9

RESULT 5
US-10-187-088-10
Sequence 10, Application US/10187088
Publication No. US20030044867A1
GENERAL INFORMATION:
APPLICANT: Pioneer HiBred International

APPLICANT: Bulla, Lee A.
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
FILE REFERENCE: 27112-20037.13
CURRENT APPLICATION NUMBER: US/10/187,088
CURRENT FILING DATE: 2002-06-28
PRIOR APPLICATION NUMBER: US/09/457,864
PRIOR FILING DATE: 1999-12-10
PRIOR APPLICATION NUMBER: US 08/326,117
PRIOR FILING DATE: 1994-10-19
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 92
TYPE: PRT
ORGANISM: Protocadherin
US-10-187-088-10

Query Match 96.0%; Score 24; DB 9; Length 92;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
|
|
|
|
Db 84 DTNDN 88

RESULT 6
US-09-457-865-5
Sequence 5, Application US/09457865
Publication No. US2003007652A1
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/457,865
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,117
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-457-865-5

Query Match 96.0%; Score 24; DB 9; Length 93;

```
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy	1	DXNDN	5
Dd	85	DTNDN	85

RESULT 7

```

US-10-006-869-42
; Sequence 42, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL-
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-006-869-42

```

Query Match 96.0% Score 24 DB 9 Length 107

Query Match 80.0%;
 Best Local Similarity 80.0%;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Caps 0;

Qy 1 DXNDN 5
| | |
Db 99 DTNDN 103

RESULT 8

```

US-10-001-883-113
; Sequence 113, Application US/10001883
; Publication No. US20030022188A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Pluta, Jason
; APPLICANT: Ghosh, Malavika
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pro
; FILE REFERENCE: DEX-0271
; CURRENT APPLICATION NUMBER: US/10/001,883
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,059
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 113
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-883-113

```

Query Match	96.0%	Score 24:	DB 9:	Length 155:
-------------	-------	-----------	-------	-------------

```
Query Match: 100.0%; Score: 24; Length: 135;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 DXNDN 5
p/p 116 DTNDN 120

RESULT 9

Query Match 96.0%; Score 24; DB 9; Length 165;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4: Conservative 0: Mismatches 1: Indels

Qy 1 DXNDN 5
94 DANDN 98
Db

RESULT 10

```

US-09-764-870-483
; Sequence 483, Application US/09764870
; Patent No. US2002042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 483
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring
; NAME/KEY: SITE

```

LOCATION: (9)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (59)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (61)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (155)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-483

Query Match 96.0%; Score 24; DB 10; Length 165;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 94 DANDN 98

RESULT 11
US-09-791-279-148
Sequence 148, Application US/09791279
Publication No. US20030050456A1
GENERAL INFORMATION:
APPLICANT: Vogel, Gabriel
APPLICANT: Wood, Linda S.
APPLICANT: Parodi, Luis
APPLICANT: Lind, Peter
TITLE OF INVENTION: No. US20030050456A1el G Protein-Coupled Receptors
FILE REFERENCE: 0048. US1

CURRENT APPLICATION NUMBER: US/09/791.279

CURRENT FILING DATE: 2001-02-23
PRIOR FILING DATE: 60/184,715
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184725
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,712
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,606
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,602
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,604
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,822
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,710
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,689
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,690
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/184,716
PRIOR FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 220
SOFTWARE: PatentIn version 3.0
SEQ ID NO 148
LENGTH: 177
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-279-148

Query Match 96.0%; Score 24; DB 9; Length 177;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5

DB 165 DSNDN 169

RESULT 12
US-09-738-626-5604
Sequence 5604, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5604
LENGTH: 212
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5604

Query Match 96.0%; Score 24; DB 9; Length 212;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 207 DTNDN 211

RESULT 13
US-10-125-540-458
Sequence 458, Application US/10125540
Publication No. US20030059875A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT214C1
CURRENT APPLICATION NUMBER: US/10/125,540
CURRENT FILING DATE: 2002-04-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 646
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 458
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (228)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (242)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature

LOCATION: (244)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (251)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (257)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-458

Query Match 96.0%; Score 24; DB 9; Length 339;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 58 DTNDN 62

RESULT 14

US-09-764-870-458
Sequence 458, Application US/09764870
Patent No. US20020042386A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT214
CURRENT APPLICATION NUMBER: US/09/764,870
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 646
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 458
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (228)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (242)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (244)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (251)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (257)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-458

Query Match 96.0%; Score 24; DB 10; Length 339;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 58 DTNDN 62

RESULT 15

US-10-125-540-482
Sequence 482, Application US/10125540
Publication No. US20030059875A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT214C1

CURRENT APPLICATION NUMBER: US/10/125,540
CURRENT FILING DATE: 2002-04-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 646
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 482
LENGTH: 427
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (278)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-482

Query Match 96.0%; Score 24; DB 9; Length 427;
Best Local Similarity 80.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 53 DTNDN 57

RESULT 16

US-09-764-870-482
Sequence 482, Application US/09764870
Patent No. US20020042386A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT214
CURRENT APPLICATION NUMBER: US/09/764,870
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 646
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 482
LENGTH: 427
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (278)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-482

Query Match 96.0%; Score 24; DB 10; Length 427;
Best Local Similarity 80.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 53 DTNDN 57

RESULT 17

US-10-125-540-296
Sequence 296, Application US/10125540
Publication No. US20030059875A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT214C1
CURRENT APPLICATION NUMBER: US/10/125,540
CURRENT FILING DATE: 2002-04-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 646
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 296
LENGTH: 453
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc feature
LOCATION: (393)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (395)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (399)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (401)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (403)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (405)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (410)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (416)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (432)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-296

Query Match 96.0%; Score 24; DB 9; Length 453;
Best Local Similarity 80.0%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 58 DTNDN 62

RESULT 18
US-09-764-870-296
Sequence 296 Application US/09764870
Patent No. US20020042386A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ14
CURRENT APPLICATION NUMBER: US/09/764,870
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 646
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 296
LENGTH: 453
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (393)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (395)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (399)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (401)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (403)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (405)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (410)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (416)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (432)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-296

Query Match 96.0%; Score 24; DB 10; Length 453;

Best Local Similarity 80.0%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 58 DTNDN 62

RESULT 19
US-09-999-162-12
Sequence 12 Application US/09999162
Patent No. US20020137171A1
GENERAL INFORMATION:
APPLICANT: Short, Jay
APPLICANT: Mathur, Eric J.
APPLICANT: Baumann, Markus
APPLICANT: Bornscheuer, Uwe T.
TITLE OF INVENTION: Hydrolase Enzymes and Their Use in Kinetic Resolution
FILE REFERENCE: DIV-011 US/112766.126
CURRENT APPLICATION NUMBER: US/09/999,162
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER: US 60/250,976
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 526
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: polypeptide 094 with hydrolase activity
US-09-999-162-12

Query Match 96.0%; Score 24; DB 10; Length 526;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 443 DANDN 447

RESULT 20
US-09-908-193-28
Sequence 28 Application US/09908193
Publication No. US20020192748A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
APPLICANT: SHIMKETS, RICHARD A.
APPLICANT: ZERHUSEN, BRYAN
APPLICANT: MALYANKAR, URIEL M.

APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-908-193-28
Query Match 96.0%; Score 24; DB 9; Length 586;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DXNDN 5
Db 236 DSDN 240
RESULT 21
US-09-893-519A-72
; Sequence 72, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Tamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAQ, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/IG548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: misc_feature

OTHER INFORMATION: Corresponds to SEQ ID NO: 145
US-09-893-519A-72
Query Match 96.0%; Score 24; DB 9; Length 587;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DXNDN 5
Db 169 DSDN 173
RESULT 22
US-09-972-086-2
; Sequence 2, Application US/09972086
; Patent No. US20020090710A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Curtis, Rory
; TITLE OF INVENTION: 57800, A NOVEL HUMAN CADHERIN AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 38155-20038.00
; CURRENT APPLICATION NUMBER: US/09/972,086
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/237,698
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-086-2
Query Match 96.0%; Score 24; DB 10; Length 589;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DXNDN 5
Db 194 DTNDN 198
RESULT 23
US-09-908-193-4
; Sequence 4, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 12.7273 Seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-8

Perfect score: 25

Sequence: 1 DXNDN 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : PIR 73:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	96.0	43	2 S35379	cadherin-related p
2	24	96.0	89	2 T17661	hypothetical prote
3	24	96.0	105	2 T25428	hypothetical prote
4	24	96.0	130	2 T22920	hypothetical prote
5	24	96.0	139	2 E97709	hypothetical prote
6	24	96.0	152	2 G90071	conserved hypotet
7	24	96.0	153	2 T06182	reverse transcript
8	24	96.0	201	2 T49213	hypothetical prote
9	24	96.0	243	2 A96744	hypothetical prote
10	24	96.0	280	2 D97273	HAD superfamily hy
11	24	96.0	295	2 S61039	hypothetical prote
12	24	96.0	296	2 T31582	hypothetical prote
13	24	96.0	313	2 E69580	arabinan-endo 1,5-
14	24	96.0	317	2 D69258	conserved hypotet
15	24	96.0	324	2 T50249	conserved flagellin
16	24	96.0	334	2 A71169	probable flagellin
17	24	96.0	338	2 S50659	hypothetical prote
18	24	96.0	341	2 S66889	probable membrane
19	24	96.0	351	2 C82808	lipoprotein precur
20	24	96.0	357	2 T33237	hypothetical prote
21	24	96.0	367	2 I39649	lipoprotein - Acti
22	24	96.0	389	2 G82140	conserved hypotet
23	24	96.0	394	1 A42115	ribosomal protein
24	24	96.0	406	2 S48763	DNA repair protein
25	24	96.0	440	2 T70703	probable uspc prot
26	24	96.0	444	2 T25413	hypothetical prote
27	24	96.0	469	2 B84644	probable protein k
28	24	96.0	489	2 A10629	probable secreted
29	24	96.0	497	2 S57073	probable membrane

30	24	96.0	508	2 T22836	hypothetical prote
31	24	96.0	529	2 S49120	laccase (EC 1.10.3
32	24	96.0	570	2 S60426	probable membrane
33	24	96.0	612	2 T39666	WD-repeat protein
34	24	96.0	620	2 T15273	hypothetical prote
35	24	96.0	621	2 T51722	transcription regu
36	24	96.0	635	2 T50661	receptor-type prot
37	24	96.0	640	2 S47865	regulatory protein
38	24	96.0	661	2 T15073	hypothetical prote
39	24	96.0	674	2 T01309	probable serine/ch
40	24	96.0	674	2 S78570	probable membrane
41	24	96.0	675	2 T02537	probable serine/ch
42	24	96.0	694	2 S68442	Grb2-associated bi
43	24	96.0	701	2 T17243	hypothetical prote
44	24	96.0	716	2 E82016	probable integral
45	24	96.0	738	2 S37876	glutamine-rich pro
46	24	96.0	742	2 S55098	probable membrane
47	24	96.0	758	2 T25751	hypothetical prote
48	24	96.0	817	2 S53919	hypothetical prote
49	24	96.0	818	2 S64251	SAP4 protein - yea
50	24	96.0	822	1 IJMSCP	P-cadherin precurs
51	24	96.0	822	2 F68812	phosphoketolase [i
52	24	96.0	826	2 B55363	desmocollin, type
53	24	96.0	829	1 IJHUCP	cadherin 3 precurs
54	24	96.0	847	2 H91155	nitrite reductase
55	24	96.0	847	2 E86001	nitrite reductase
56	24	96.0	847	2 AG1001	nitrite reductase
57	24	96.0	847	2 H65130	nitrite reductase
58	24	96.0	848	2 AF0020	cadherin - African
59	24	96.0	871	2 S47518	probable membrane
60	24	96.0	880	2 S51473	cadherin 1 precurs
61	24	96.0	882	1 IJHUCE	E-cadherin precurs
62	24	96.0	887	1 IJCHCL	protocadherin 68 -
63	24	96.0	889	2 T09055	protocadherin prec
64	24	96.0	895	1 IJXLCF	EP-cadherin precu
65	24	96.0	896	2 I45858	desmocollin - bovi
66	24	96.0	896	2 A55363	desmocollin, type
67	24	96.0	905	2 S43064	cadherin - African
68	24	96.0	961	2 S67568	probable membrane
69	24	96.0	962	2 J65808	G protein-coupled
70	24	96.0	997	2 D88473	protein F40H6.2 [i
71	24	96.0	1064	2 T44419	carbamoyl-phosphat
72	24	96.0	1065	1 I40169	carbamoyl-phosphat
73	24	96.0	1069	2 T00043	BH-protocadherin-a
74	24	96.0	1072	2 T37876	probable (Xeroderm
75	24	96.0	1072	2 T00041	BH-protocadherin P

ALIGNMENTS

RESULT 1

S35379
C:Species: Rattus norvegicus (Norway rat)
C>Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
C:Accession: S35379
R:Sano, K.; Tanihara, H.; Heimark, R.L.; Obata, S.; Davidson, M.; St John, T.; Taketa
EMBO J. 12, 2249-2256, 1993
A:Title: Protocadherins: a large family of cadherin-related molecules in central nerv
A:Reference number: S35379; MUID:93285094; PMID:8508762
A:Accession: S35379
A:Molecule type: mRNA
A:Residues: 1-43 <SAN>

Query Match 96.0%; Score 24; DB 2; Length 43;
Best Local Similarity 80.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DXNDN 5
| | | | |
DB 37 DANDN 41

RESULT 2

T17661
hypothetical protein al70L - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17661
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17661
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-89 <GRA>
A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96538.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: al70L

Query Match 96.0%; Score 24; DB 2; Length 89;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
DB 43 DSNDN 47

RESULT 3

T25428
hypothetical protein T28F4.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25428
R;McMurray, A.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z20033
A;Accession: T25428
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-105 <WIL>
A;Cross-references: EMBL:Z72517; PIDN:CAA96692.1; GSPDB:GN00019; CESP:T28F4.5
A;Experimental source: clone T28F4
C;Genetics:
A;Gene: CESP:T28F4.5
A;Map position: 1
A;Introns: 61/3

Query Match 96.0%; Score 24; DB 2; Length 105;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
DB 36 DSNDN 40

RESULT 4

T22920
hypothetical protein F58E6.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22920
R;McMurray, A.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19638
A;Accession: T22920
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-130 <WIL>
A;Cross-references: EMBL:Z70754; PIDN:CAA94774.1; GSPDB:GN00023; CESP:F58E6.4
A;Experimental source: clone F58E6
C;Genetics:

A;Gene: CESP:F58E6.4
A;Map position: 5
A;Introns: 91/3

Query Match 96.0%; Score 24; DB 2; Length 130;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
DB 94 DTNDN 98

RESULT 5

E97709
hypothetical protein RC0077 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 10-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C;Accession: E97709
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893

A;Accession: E97709
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-139 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL02615.1; PID:gi5619115; GSPDB:GN00173
C;Genetics:
A;Gene: RC0077
C;Superfamily: Rickettsia prowazekii hypothetical protein RP051

Query Match 96.0%; Score 24; DB 2; Length 139;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
DB 129 DSNDN 133

RESULT 6

G90071
conserved hypothetical protein SA2432 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: G90071
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Cui, A.; Mizukami-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: G90071
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-152 <KUR>
A;Cross-references: GB:BA000018; PID:gi3702597; PIDN:BA843737.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA2432

Query Match 96.0%; Score 24; DB 2; Length 152;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
DB 71 DTNDN 75

RESULT 7

T06182

reverse transcriptase homolog - barley
C;Species: Hordeum vulgare [barley]
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
C;Accession: T06182
R;Liu, K.; Somerville, S.
Genome 39, 1159-1168, 1996
A;Title: Cloning and characterization of a highly repeated DNA sequence in Hordeum vulgare
A;Reference number: 215514; MUID:97137848; PMID:8983185
A;Accession: T06182
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-153 <LIU>
A;Cross-references: EMBL:L23524; NID:g387901; PIDN:AAB42154.1; PID:g387902
A;Experimental source: cv.CI-16137

Query Match 96.0%; Score 24; DB 2; Length 153;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
Db 66 DANDN 70

RESULT 8
T49213
hypothetical protein F27K19.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C;Accession: T49213
R;Benes, V.; Wurmbsch, E.; Drzonek, H.; Ansoorge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000
A;Reference number: 225014
A;Accession: T49213
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <BEN>
A;Cross-references: EMBL:AL163832; GSPDB:GN00061; ATSP:F27K19.190
A;Experimental source: cultivar Columbia; BAC clone F27K19
C;Genetics:
A;Gene: ATSP:F27K19.190
A;Map position: 3
A;Introns: 12/3
C;Superfamily: Arabidopsis thaliana hypothetical protein F27K19.190

Query Match 96.0%; Score 24; DB 2; Length 201;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
Db 53 DTNDN 57

RESULT 9
A96744
hypothetical protein F28P5.14 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A96744
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A96141; MUID:21016719; PMID:11130712
A;Accession: A96744
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-243 <STO>
A;Cross-references: GB:AE005173; NID:g11054408; PIDN:AAG27795.1; GSPDB:GN00141
C;Genetics:
A;Gene: P28P5.14
A;Map position: 1

Query Match 96.0%; Score 24; DB 2; Length 243;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
Db 191 DANDN 195

RESULT 10
D97273
HAD superfamily hydrolase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97273
R;Nolling, J.; Breton, G.; Omselchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.;
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97273
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80975.1; PID:g15026094; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3035
C;Superfamily: hypothetical protein ywpJ

Query Match 96.0%; Score 24; DB 2; Length 280;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
Db 260 DTNDN 264

RESULT 11
S61039
hypothetical protein YDL173w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein D1438
C;Species: Saccharomyces cerevisiae
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
C;Accession: S61039; S67725
R;Pohl, T.M.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61010
A;Accession: S61039
A;Molecule type: DNA
A;Residues: 1-295 <POH>
A;Cross-references: EMBL:Z67750; NID:g1061256; PIDN:CAA91572.1; PID:g1061265
R;Pohl, T.M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67708
A;Accession: S67725
A;Molecule type: DNA
A;Residues: 1-295 <POH>
A;Cross-references: EMBL:Z74221; NID:g1431275; PIDN:CAA98746.1; PID:e253081; PID:g14
C;Genetics:
A;Experimental source: strain S288C
A;Cross-references: SGD:S0002332
A;Map position: 4L

Query Match 96.0%; Score 24; DB 2; Length 295;

Best Local Similarity 80.0%; Pred. No. 3.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0

QY 1 DXNDN 5
|
|
|
Db 92 DSNDN 96

RESULT 12

T31582

hypothetical protein Y105CSA.Y - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31582

R:McMurray, A.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z21045

A:Accession: T31582

A:Status: Preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-236 <WIL>

A:Cross-references: EMBL:AL117193; NID:e1549703; PIDN:CAB55005.1; CESP:Y105CSA.Y

A:Experimental source: clone Y105CSA

C:Genetics:

A:Gene: CESP:Y105CSA.Y

A:Introns: 22/1; 75/1

Query Match 96.0%; Score 24; DB 2; Length 296;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
Db 83 DSNDN 87

RESULT 13

E69580

arabinan-endo 1,5-alpha-L-arabinase abnA - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: E69580

R:Kunat, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, H.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: E69580

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-313 <KUN>

A:Cross-references: GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CAB14841.1; PID:e1184130;

A:Experimental source: strain 168

C:Genetics:

A:Gene: abnA

Query Match 96.0%; Score 24; DB 2; Length 313;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|

Db 300 DANDN 304

RESULT 14

D69258

conserved hypothetical protein AF0068 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: D69258

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,

C:Accession: D69258

J:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arc

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69258

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-317 <KLE>

A:Cross-references: GB:AB001102; GB:AE000782; NID:g2689425; PIDN:AB91161.1; PID:g26

Query Match 96.0%; Score 24; DB 2; Length 317;

Best Local Similarity 80.0%; Pred. No. 3.4e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
Db 212 DANDN 216

RESULT 15

T50249

conserved hypothetical protein SPAC694.04c [imported] - fission yeast (Schizosacchar

C:Species: Schizosaccharomyces pombe

C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000

C:Accession: T50249

R:Art, R.; Volckaert, G.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, February 2000

A:Reference number: Z25049

A:Accession: T50249

A:Status: Preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-324 <ASR>

A:Cross-references: EMBL:AL138666; PIDN:CAB71842.1; GSPDB:GN00066; SPDB:SPAC694.04c

A:Experimental source: strain 972h(-); cosmid c694

C:Genetics:

A:Gene: SPDB:SPAC694.04c

A:Map position: 1

A:Introns: 19/1

C:Superfamily: Arabidopsis thaliana hypothetical protein F2K15.180

Query Match 96.0%; Score 24; DB 2; Length 324;

Best Local Similarity 80.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
Db 128 DANDN 132

RESULT 16

A71169

probable flagellin B precursor - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000

C:Accession: A71169

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; S

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Og

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: A71169
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-334 <KAW>
A:Cross-references: GB:AP000002; NID:g3236129; PIDN:BAA29638.1; PID:g3256955
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0549
C:Superfamily: archaeal flagellin

Query Match 96.0%; Score 24; DB 2; Length 334;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
|
|
|
Db 171 DSDND 175

RESULT 17

S50659
hypothetical protein YER156c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002
C:Accession: S50659
R:Dieterich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmids 8229, 9115, 9132, 9981, and lambda
A:Reference number: S50430
A:Accession: S50659
A:Molecule type: DNA
A:Residues: 1-338 <DIE>
A:Cross-references: EMBL:U18917; NID:g603377; PID:g603396; GSPDB:GN000005; MIPS:YER156C
C:Genetics:
A:Gene: MIPS:YER156C
A:Cross-references: SGD:S0000958
A:Map position: 5R
C:Superfamily: Arabidopsis thaliana hypothetical protein F2K15.180

Query Match 96.0%; Score 24; DB 2; Length 338;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
|
|
|
Db 137 DANDN 141

RESULT 18

S66689
probable membrane protein YOL007c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O2373
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S66689
R:Hughes, B.; Fohl, T.M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66685
A:Accession: S66689
A:Molecule type: DNA
A:Residues: 1-341 <HUG>
A:Cross-references: EMBL:Z74749; NID:g1419772; PID:g251834; PID:g1419773; GSPDB:GN000015;
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:CSI2; MIPS:YOL007c
A:Cross-references: SGD:S0005367
A:Map position: 15L
C:Keywords: transmembrane protein
F:11-27/Domain: transmembrane #status predicted <TM1>
F:133-145/Domain: transmembrane #status predicted <TM2>

Query Match 96.0%; Score 24; DB 2; Length 341;

Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
|
|
|
Db 235 DANDN 239

RESULT 19

C82808
lipoprotein precursor XF0416 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: C82808
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Se
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20385717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82808
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <SIM>
A:Cross-references: GB:AE003893; GB:AE003849; NID:g9105253; PIDN:AAF83226.1; GSPDB:G
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; F
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; L
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martin
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Saw
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Sil
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0416

Query Match 96.0%; Score 24; DB 2; Length 351;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
|
|
|
Db 319 DTNDN 323

RESULT 20

T33237
hypothetical protein T10H9.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33237
R:Greco, T.; Bradshaw, H.; O'Brien, D.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid T10H9.
A:Reference number: Z21306
A:Accession: T33237
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <GRE>
A:Cross-references: EMBL:AF067949; PIDN:AAC19235.1; GSPDB:GN000023; CESP:T10H9.1
A:Experimental source: strain Bristol N2; clone T10H9
C:Genetics:
A:Gene: CESP:T10H9.1
A:Map position: 5
A:Introns: 17/3; 70/1; 135/3; 163/3; 225/3; 293/2

Query Match 96.0%; Score 24; DB 2; Length 357;

Best Local Similarity 80.0%; Pred. No. 3.8e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 4; Conservative 0;

QY 1 DXNDN 5
|
|
|
|
Db 347 DSDND 351

RESULT 21

I39649
C:Species: Actinobacillus pleuropneumoniae
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 17-Nov-2000
C:Accession: I39649; I39518; S52740
R:Bunka, S.; Christensen, C.; Potter, A.A.; Willson, P.J.; Gerlach, G.F.
Infect. Immun. 63, 2797-2800, 1995
A:Title: Cloning and characterization of a protective outer membrane lipoprotein of Actinobacillus pleuropneumoniae
A:Reference number: I39649; MUID:95310047; PMID:7790104
A:Accession: I39649
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-367 <RES>
A:Cross-references: EMBL:248920; NID:G757528; PIDN:CAA88756.1; PID:G757529
R:Ito, H.; Uchida, I.; Sekizaki, T.; Ootshi, E.; Kawai, T.; Taneno, A.; Terak
Microb. Pathog. 18, 29-36, 1995
A:Title: Molecular cloning of an Actinobacillus pleuropneumoniae outer membrane lipoprotein
A:Reference number: I39518; MUID:95302932; PMID:7783596
A:Accession: I39518
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-361, 'KEDK', 366 <RES>
A:Cross-references: GB:D28491; NID:G808858; PIDN:BA005852.1; PID:G808859
C:Genetics:
A:Gene: om1a

Query Match 96.0%; Score 24; DB 2; Length 367;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
Db 151 DSDND 155

RESULT 22

G82140
C:Species: conserved hypothetical protein VC1912 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82140
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Emdoliaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82140
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-389 <HEI>
A:Cross-references: GB:AE004267; GB:AE003852; NID:G9656446; PIDN:AAF95060.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1912
A:Map position: 1
C:Superfamily: hypothetical protein H11223; tetratricopeptide repeat homology

Query Match 96.0%; Score 24; DB 2; Length 389;
Best Local Similarity 80.0%; Pred. No. 4.2e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|

Db 193 DSDND 197

RESULT 23

A42115
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 23-Mar-2001
C:Accession: A42115; S46799; S27429
R:Davis, S.C.; Tzagoloff, A.; Ellis, S.R.
J. Biol. Chem. 267, 5508-5514, 1992
A:Title: Characterization of a yeast mitochondrial ribosomal protein structurally related to the human mitochondrial ribosomal protein S2
A:Reference number: A42115; MUID:92184810; PMID:1531984
A:Accession: A42115
A:Molecule type: DNA
A:Residues: 1-394 <DAV>
A:Cross-references: EMBL:M82841; NID:G171981; PIDN:AAA34793.1; PID:G171982
A:Note: sequence extracted from NCBI backbone (NCBIN:87499, NCBIP:87502)
R:Favella, T.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9780.
A:Reference number: S46799
A:Accession: S46799
A:Molecule type: DNA
A:Residues: 1-394 <FAV>
A:Cross-references: EMBL:U10555; NID:G500813; PIDN:AAB68428.1; PID:G500819; GSPDB:G1
C:Genetics:
A:Gene: SGD:MRP4; MIPS:YHL004W
A:Cross-references: SGD:S0000996; MIPS:YHL004W
A:Map position: 8L
C:Superfamily: yeast ribosomal protein S2
C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 96.0%; Score 24; DB 1; Length 394;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
|
|
|
|
Db 296 DSDND 300

RESULT 24

S48763
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Oct-1999
C:Accession: S48763; S48831; S48889; S55823; S67893
R:Coster, F.; Jonniaux, J.L.; Goffeau, A.
submitted to the EMBL Data Library, October 1994
A:Reference number: S48758
A:Accession: S48763
A:Molecule type: DNA
A:Residues: 1-406 <COS>
A:Cross-references: EMBL:X82086; NID:G558241; PIDN:CAA57603.1; PID:G558247
R:Richards, C.; Harris, D.E.
submitted to the EMBL Data Library, November 1994
A:Reference number: S49823
A:Accession: S49831
A:Molecule type: DNA
A:Residues: 1-406 <RIC>
A:Cross-references: EMBL:Z46796; NID:G577794; PIDN:CAA86798.1; PID:G577803
R:Lovett, S.T.
Gene 142, 103-106, 1994
A:Title: Sequence of the RAD55 gene of Saccharomyces cerevisiae: similarity of RAD55
A:Reference number: S48889; MUID:94237473; PMID:8181742
A:Accession: S48889
A:Molecule type: DNA
A:Residues: 1-109, 'L', 111-196, 'N', 198-406 <LOV>
A:Cross-references: EMBL:U01144; NID:G436820; PIDN:AAA19688.1; PID:G509494
R:Coster, F.; Jonniaux, J.L.; Goffeau, A.

Yeast 11, 673-679, 1995
 A;Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr
 A;Reference number: S55819; MUID:96093910; PMID:7483840
 A;Accession: S55823
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-406 <COW>
 A;Cross-references: EMBL:X82086; NID:g558241; PIDN:CAA57603.1; PID:g558247
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
 R;Pouy, F.; Jonniaux, J.L.; Purnelle, B.; Coster, F.; Goffeau, A.
 submitted to the Protein Sequence Database, July 1996
 A;Reference number: S67889
 A;Accession: S67893
 A;Molecule type: DNA
 A;Residues: 1-406 <FOU>
 A;Cross-references: EMBL:Z74372; NID:gl431540; PIDN:CAA98895.1; PID:e2533328; PID:gl43154
 A;Experimental source: strain S288C
 C;Genetics:
 A;Gene: SGD:RAD55
 A;Cross-references: SGD:S0002483; MIPS:YDR076w
 A;Map position: 4R
 C;Function:
 A;Description: involved in recombination and recombinational DNA repair
 C;Keywords: DNA repair

Query Match 96.0%; Score 24; DB 2; Length 406;
 Best Local Similarity 80.0%; Pred. NO. 4.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 ||||
 Db 385 DSNDN 389

RESULT 25

G70703
 Probable uspC protein, Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C;Accession: G70703
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: G70703
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-440 <COL>
 A;Cross-references: GB:Z79702; GB:AL123456; NID:g3261642; PIDN:CAB02072.1; PID:e264240;
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: uspC

Query Match 96.0%; Score 24; DB 2; Length 440;
 Best Local Similarity 80.0%; Pred. NO. 4.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 ||||
 Db 267 DTNDN 271

Search completed: June 20, 2003, 20:58:51
 Job time : 19.7273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 ; Search time 6.02273 Seconds
(without alignments)
34.433.Million cell updates/sec

Title: US-10-105-008-8

Perfect score: 25

Sequence: 1 DXNDN 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	96.0	324	YI4 SCHPO	Q9p7t6 schizosacch
2	24	96.0	334	FLA3 PYRHO	O58284 pyrococcus
3	24	96.0	338	YEY6 YEAST	P40093 saccharomyc
4	24	96.0	341	CS12 YEAST	Q08054 saccharomyc
5	24	96.0	341	MBNL_MOUSE	Q9jke5 mus musculu
6	24	96.0	388	MBNL_HUMAN	Q9nr56 homo sapien
7	24	96.0	394	RT04 YEAST	P32902 saccharomyc
8	24	96.0	406	RA55 YEAST	P38953 saccharomyc
9	24	96.0	422	TKSU PYRKO	P58502 pyrococcus
10	24	96.0	433	PEPD_SALDU	O85304 salmonella
11	24	96.0	497	XJ24 YEAST	P47114 saccharomyc
12	24	96.0	529	LAC1_PLEOS	Q12729 pleurotus o
13	24	96.0	570	YGR0 YEAST	P53109 saccharomyc
14	24	96.0	640	SP10 YEAST	P35208 saccharomyc
15	24	96.0	674	YFE2 YEAST	P43560 saccharomyc
16	24	96.0	738	YKF4 YEAST	P35732 saccharomyc
17	24	96.0	742	KM65 YEAST	Q03656 saccharomyc
18	24	96.0	758	YLDM_CABEL	Q94255 caenorhabdi
19	24	96.0	776	CDBF_HUMAN	Q9nr17 homo sapien
20	24	96.0	787	CDBF_HUMAN	Q9y5e8 homo sapien
21	24	96.0	788	CADA_HUMAN	Q9y6n8 homo sapien
22	24	96.0	793	CDB7_HUMAN	Q9y5e2 homo sapien
23	24	96.0	794	CDB6_HUMAN	Q9y5e3 homo sapien
24	24	96.0	795	CDB4_HUMAN	Q9y5e5 homo sapien
25	24	96.0	795	CDB5_HUMAN	Q9y5e4 homo sapien
26	24	96.0	795	CDBC_HUMAN	Q9y5f1 homo sapien
27	24	96.0	796	CDB3_HUMAN	Q9y5e6 homo sapien
28	24	96.0	797	CDB9_HUMAN	Q9y5e1 homo sapien
29	24	96.0	797	CDBB_HUMAN	Q9y5f2 homo sapien
30	24	96.0	798	CDB2_HUMAN	Q9y5e7 homo sapien
31	24	96.0	798	CDBD_HUMAN	Q9y5f0 homo sapien
32	24	96.0	798	CDBE_HUMAN	Q9y5e9 homo sapien
33	24	96.0	800	CDBA_HUMAN	Q9un67 homo sapien

34	24	96.0	801	CDB8_HUMAN	Q9un66 homo sapien
35	24	96.0	817	YGA4 YEAST	P46949 saccharomyc
36	24	96.0	818	SAP4 YEAST	P53036 saccharomyc
37	24	96.0	822	CAD3_MOUSE	P10287 mus musculu
38	24	96.0	822	PHK_LACLA	Q9cfh4 lactococcus
39	24	96.0	829	CAD3_HUMAN	P22223 homo sapien
40	24	96.0	847	NIRE_ECOLI	P08201 escherichia
41	24	96.0	850	D7_DICDI	P54692 dictyosteli
42	24	96.0	872	CAD1_XENLA	P30944 xenopus lae
43	24	96.0	880	CAD1_XENLA	P33148 xenopus lae
44	24	96.0	880	YL86 YEAST	Q06708 saccharomyc
45	24	96.0	882	CAD1_HUMAN	P12830 homo sapien
46	24	96.0	884	CAD6_XENLA	P33152 xenopus lae
47	24	96.0	887	CAD1_CHICK	P08641 gallus gall
48	24	96.0	895	DSC3_MOUSE	P5850 mus musculu
49	24	96.0	896	DSC3_BOVIN	Q28050 bos taurus
50	24	96.0	896	DSC3_HUMAN	Q14574 homo sapien
51	24	96.0	936	CDA5_HUMAN	Q9y5h7 homo sapien
52	24	96.0	947	CDA4_HUMAN	Q9un74 homo sapien
53	24	96.0	948	CDA4_HUMAN	Q9y5i2 homo sapien
54	24	96.0	949	CDA8_HUMAN	Q9y5i1 homo sapien
55	24	96.0	963	CHC1_HUMAN	Q9h158 homo sapien
56	24	96.0	997	YPX2_CABEL	Q20256 caenorhabdi
57	24	96.0	1007	CHC2_HUMAN	Q9y5i4 homo sapien
58	24	96.0	1064	CARB_BACST	O50302 bacillus st
59	24	96.0	1065	CARB_BACCL	P46537 bacillus ca
60	24	96.0	1069	PCD7_HUMAN	O60245 homo sapien
61	24	96.0	1097	S24C_ARATH	Q9m291 arabidopsis
62	24	96.0	1180	PC12_MOUSE	O55134 mus musculu
63	24	96.0	1184	PC12_HUMAN	Q9npg4 homo sapien
64	24	96.0	1220	PTC1_BRARE	Q98864 brachydanio
65	24	96.0	1351	RPOM YEAST	P13433 saccharomyc
66	24	96.0	1858	VGNE_CPSMV	P36312 cowpea seve
67	24	96.0	1943	PC15_MOUSE	Q99p11 mus musculu
68	24	96.0	1955	CLR1_HUMAN	Q96qu1 homo sapien
69	24	96.0	3014	CLR1_MOUSE	Q9nyq6 homo sapien
70	24	96.0	3034	CLR1_MOUSE	O35161 mus musculu
71	24	96.0	3298	PC16_HUMAN	Q96iq0 homo sapien
72	24	96.0	3411	POLG_YEFV1	P03314 y genome po
73	24	96.0	3411	POLG_YEFV2	P19901 y genome po
74	24	96.0	3579	STAN_DROME	Q9vsn8 drosophila
75	24	96.0	4349	FAT2_HUMAN	Q9nyq8 homo sapien

ALIGNMENTS

RESULT 1
YI4 SCHPO
ID YI4 SCHPO STANDARD; PRT; 324 AA.
AC Q9p7t6
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C694.04c in chromosome 1.
GN SPAC694.04C
OS Schizosaccharomyceti pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
RN [1]_NCBI_TaxID=4896;
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RA MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Piat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsals K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Sammonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Lucas M., Rochet M., Gallardin C., Talla V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE UPF0160 (MYG1) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL138666; CAB71842.1;
 DR InterPro; IPR003226; Me_prt_hydrlase.
 DR Pfam; PF03690; UPF0160; 1.
 DR ProDom; PD034736; Me_prt_hydrlase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 324 AA; 36672 MW; 2E8D25E2BAB116FA CRC64;
 Query Match 96.0%; Score 24; DB 1; Length 324;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 128 DANDN 132
 RESULT 2
 FLA3 PYRHO STANDARD; PRT; 334 AA.
 AC O58284;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable flagellin PH0549.
 GN PH0549.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Pyrococcus.
 CC NCBI TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3.
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3.";
 RL DNA Res. 5:55-76(1998).
 CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
 CC FORM THE FILAMENTS OF FLAGELLA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEL FLAGELLIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP000002; BAA29638.1;
 DR InterPro; IPR002774; Arch_flagellin.
 DR Pfam; PF01917; Arch_flagellin; 1.
 KW Flagella; Multigene family; Complete proteome.
 SQ SEQUENCE 334 AA; 35438 MW; 9CEAAA6E07C92296 CRC64;
 Query Match 96.0%; Score 24; DB 1; Length 334;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 171 DSNDN 175
 RESULT 3
 YE16_YEAST STANDARD; PRT; 338 AA.
 AC P40093;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 38.2 kDa protein in BEM2-NCBI intergenic region.
 GN YER156C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC NCBI TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D., Oh C.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE UPF0160 (MYG1) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U18917; AAB64683.1;
 DR SGD; S0000958; YER156C.
 DR InterPro; IPR003226; Me_prt_hydrlase.
 DR Pfam; PF03690; UPF0160; 1.
 DR ProDom; PD034736; Me_prt_hydrlase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 338 AA; 38174 MW; 5B13FC27EFE911FF CRC64;
 Query Match 96.0%; Score 24; DB 1; Length 338;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXNDN 5
 DB 137 DANDN 141

RESULT 4
CS12_YEAST STANDARD; PRT; 341 AA.
ID CS12_YEAST
AC Q08054;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Chitin synthase 3 complex protein CS12.
GN CS12 OR YOL007C.
OS Saccharomyces cerevisiae (Baker's yeast)
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Hughes B., Pohl T.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION.
RA Demarini D.J., Fringle J.R.;
RL Unpublished observations (FEB-1998).
CC -1- FUNCTION: APPEARS TO BE A STRUCTURAL COMPONENT OF THE CHITIN
CC SYNTHASE 3 COMPLEX.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z74749; CAA99006.1; --
DR SGI; S0005367; CS12.
FT DOMAIN 33 112 POLY-SER.
SQ SEQUENCE 341 AA; 37256 MW; 5F9CCD1283890FP9 CRC64;
Query Match 96.0%; Score 24; DB 1; Length 341;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 DXNDN 5
DB 235 DANDN 239
RESULT 5
MBNL_MOUSE STANDARD; PRT; 341 AA.
ID MBNL_MOUSE
AC Q9JKF5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Muscleblind-like protein (Triplet-expansion RNA-binding protein).
GN MBNL OR EXP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Castagnola P., Monticone M., Boreani G., Bassi M.T., Tonachini L.;
RL "cDNA cloning of mouse muscleblind."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Binds to CUG triplet repeat expansion dsRNA (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: CONTAINS 4 C3H1-TYPE ZINC FINGERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF231110; AAF72159.1; --
DR MGD; MGI:1928482; Mbnl.
DR InterPro: IPR000571; Znf.CCCH.
DR Pfam: PF00642; zf.CCCH; 4.
DR SMART; SM00356; Znf.C3H1; 4.
KW Zinc-finger; Repeat; Nuclear protein; RNA-binding.
FT ZN_FING 14 40 C3H1-TYPE 1.
FT ZN_FING 47 72 C3H1-TYPE 2.
FT ZN_FING 178 205 C3H1-TYPE 3.
FT ZN_FING 215 239 C3H1-TYPE 4.
SQ SEQUENCE 341 AA; 36975 MW; 8E008DB5C7EF8AB9 CRC64;
Query Match 96.0%; Score 24; DB 1; Length 341;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 DXNDN 5
DB 211 DTNDN 215
RESULT 6
MBNL_HUMAN STANDARD; PRT; 388 AA.
ID MBNL_HUMAN
AC Q5NR56; O43311; O43797;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Muscleblind-like protein (Triplet-expansion RNA-binding protein).
GN MBNL OR EXP OR KIAA0428.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Borsani G., Barbieri A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=98116655; PubMed=9455477;
RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O., et al.
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro."
RN DNA Res. 4:307-313(1997).
RP SEQUENCE FROM N.A. (ISOFORM 3), ALTERNATIVE SPLICING, AND
RP CHARACTERIZATION.
RX MEDLINE=20428414; PubMed=10970838;
RA Miller J.W., Urbanati C.R., Teng-Umuay P., Stenberg M.G., Byrne B.J.,
RA Thornton C.A., Swanson M.S.;
RT "Recruitment of human muscleblind proteins to (CUG) (n) expansions
RT associated with myotonic dystrophy."
RL EMBO J. 19:4439-4448(2000).
CC -1- FUNCTION: Binds to CUG triplet repeat expansion dsRNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; EXP42 (shown here), EXP40 and
CC EXP35; are produced by alternative splicing.
CC -1- DISEASE: Muscular dystrophy may be caused by aberrant recruitment
CC of MBNL to the DMPK transcript (CUG) (n)-expansion.
CC -1- SIMILARITY: CONTAINS 4 C3H1-TYPE ZINC FINGERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL; Y13829; CAA74155.1; --
 CC EMBL; AB007888; BAA24858.1; --
 CC EMBL; AF253334; AAF76138.1; --
 CC Genbank; HGNC:6923; MENL.
 CC MIM; 606516; --
 CC InterPro; IPR000571; Znf CCH; 6.
 CC Pfam; PF00642; zf-CCH; 6.
 CC SMART; SM00356; Znf C3H1; 6.
 CC Zinc-finger; Repeat; Nuclear protein; RNA-binding;
 KW Alternative splicing.
 FT ZN FING 14 40 C3H1-TYPE 1.
 FT ZN FING 47 72 C3H1-TYPE 2.
 FT ZN FING 179 206 C3H1-TYPE 3.
 FT ZN FING 216 240 C3H1-TYPE 4.
 FT VARSPLIC 116 183 MISSING (IN ISOFORM EXP35).
 FT VARSPLIC 270 287 MISSING (IN ISOFORM EXP35 AND ISOFORM
 EXP40).
 SQ SEQUENCE 388 AA; 41817 MW; 118D256A81A86695 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 388;
 Best Local Similarity 80.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 212 DTNDN 216

RESULT 7
 ID RT04 YEAST STANDARD; PRT; 394 AA.
 AC P32902;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Mitochondrial 40S ribosomal protein MRP4.
 GN MRP4 OR YHL004W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92184810; PubMed=1531984;
 RA Davis S.C., Tzagoloff A., Ellis S.R.;
 RT "Characterization of a yeast mitochondrial ribosomal protein
 RT structurally related to the mammalian 68-kDa high affinity laminin
 RT receptor."
 RL J. Biol. Chem. 267:5508-5514 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y., Mouser L.,
 RA Lathille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 RA Nham M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII."
 RL Science 265:2077-2082 (1994).
 CC -1- FUNCTION: COMPONENT OF THE SMALL SUBUNIT OF MITOCHONDRIAL
 CC RIBOSOME.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL; M82841; AAA34793.1; --
 CC EMBL; U10555; AAB68428.1; --
 CC PIR; S27429; S27429.
 CC PIR; A42115; A42115.
 CC SGD; S0000996; MRP4.
 CC InterPro; IPR001865; Ribosomal S2.
 CC Pfam; PF00318; Ribosomal S2; 1.
 CC PRINTS; PR00395; RIBOSOMALS2.
 CC TIGRFAMs; TIGR01011; rpsB bact; 1.
 CC PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
 CC PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
 KW Ribosomal protein; Mitochondrion.
 SQ SEQUENCE 394 AA; 44151 MW; 81A1CB33134BC559 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 394;
 Best Local Similarity 80.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 296 DSNDN 300

RESULT 8
 ID RA55 YEAST STANDARD; PRT; 406 AA.
 AC P38953;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA repair protein RAD55.
 GN RAD55 OR YDR076W OR D4426.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB320;
 RX MEDLINE=94237473; PubMed=8181742;
 RA Lovett S.T.;
 RT "Sequence of the RAD55 gene of Saccharomyces cerevisiae: similarity
 RT of RAD55 to prokaryotic RecA and other RecA-like proteins."
 RL Gene 142:103-106 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RA Kuwahara S., Yoshida T., Fong J., Ogawa T.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1679;
 RA Coster F., Jonniaux J.-L., Goffeau A.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Foury F., Jonniaux J.-L., Purnelle B., Coster F., Goffeau A.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR RADIATION RESISTANCE AND MEIOTIC VIABILITY
 CC AND PRESUMABLY ACTS IN RECOMBINATION AND RECOMBINATIONAL DNA
 CC REPAIR PATHWAYS.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: BELONGS TO THE RECA FAMILY. RAD55 SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z46796; CAA86798.1; -
EMBL; D10481; BAA01284.1; -
EMBL; U01144; AAA19688.1; -
EMBL; X82086; CAA57603.1; -
EMBL; Z74372; CAA38895.1; -
PIR; S48763; S48763.
SGD; S0002483; RAD55.
InterPro; IPR001553; RecA.
ProDom; PD000229; RecA; 1.
PROSITE; PS0162; RECA_2; 1.
KW DNA damage; DNA repair; ATP-binding; Nuclear protein.
FT NP BIND 43 50
ATP (POTENTIAL).
FT DOMAIN 277 287
SER-RICH.
FT DOMAIN 371 378
SER-RICH.
FT CONFLICT 110 110
F -> L (IN REF. 3).
FT CONFLICT 197 197
F -> N (IN REF. 3).
SQ SEQUENCE 406 AA; 46349 MW; 8EA49831F7F3E098 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 406;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DXNDN 5
DB 385 DSDND 389

RESULT 9

TKSU PYRKO STANDARD; PRT; 422 AA.
AC P58502; Q977F5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tk-subtilisin precursor (EC 3.4.21.-).
OS Pyrococcus kodakaraensis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=69014;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=KOD1;
RX MEDLINE=21268834; PubMed=11375149;
RA Kannan Y., Koga Y., Inoue Y., Haruki M., Takagi M., Imanaka T.,
Morikawa M., Kanaya S.;
RT "Active subtilisin-like protease from a hyperthermophilic archaeon in
a form with a putative prosequence."
RL Appl. Environ. Microbiol. 67:2445-2452(2001)
CC -!- FUNCTION: Has a broad substrate specificity with a slight
preference to large hydrophobic amino acid residues at the P1
position.
CC -!- COFACTOR: Calcium.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: Thermostable; high activity at 80 degrees Celsius.
CC Optimal pH is 9.5.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB056701; BAB60701.1; -
InterPro; IPR000209; Peptidase_S8.
Pfam; PF00082; Peptidase_S8; 1.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Serine protease; Calcium-binding; Zymogen; Signal.
FT SIGNAL 1 24
POTENTIAL.
FT PROPEP 25 106
POTENTIAL.
FT CHAIN 107 422
TK-SUBTILISIN.
FT ACT SITE 139 139
CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 177 177
CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 348 348
CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 422 AA; 43785 MW; 843255BCD806DB71 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 422;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 310 DSDND 314

RESULT 10

PEPD SALDU STANDARD; PRT; 433 AA.
AC O85304;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable dipeptidase (EC 3.4.-.-).
GN PIPD.
OS Salmonella dublin.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=98360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2229;
RA Wood M.W., Jones M.A., Watson P.R., Hedges S., Wallis T.S.,
Gal'vov E.E.;
RT "Identification of a pathogenicity island required for Salmonella
enteropathogenicity."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: DIPEPTIDE + H(2)O = 2 AMINO ACID.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY U34.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF050858; AAC33721.1; -
MEROPS; U34.UPW; -
InterPro; IPR005322; Peptidase U34.
Pfam; PF03577; Peptidase_U34; 1.
KW Hydrolase; Dipeptidase.
SQ SEQUENCE 433 AA; 48159 MW; 90EDF36A8100BB9C CRC64;

Query Match 96.0%; Score 24; DB 1; Length 433;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 89 DTNDN 93

RESULT 11

YJ24_YEAST STANDARD; PRT; 497 AA.
 ID_YJ24_YEAST
 AC P47114;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 57.5 kDa protein in RAD7-HIT1 intergenic region.
 GN YR054W OR J1669.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=864037976; PubMed=8840504;
 RA Huang M.-E., Manus V., Chuat J.-C., Galibert F.;
 RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
 RT reading frames and a gene cluster with a counterpart on chromosome
 RT XI.";
 RL Yeast 12:869-875(1996).
 CC -1- FUNCTION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: TO YEAST YML047C.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z49554; CAAB89582.1; -;
 CC EMBL; L47993; AAB39280.1; -;
 CC SGD; S0003815; YJR054W.
 KW Hypothetical protein: Transmembrane.
 FT TRANSMEM 30 50 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 217 237 POTENTIAL.
 SQ SEQUENCE 497 AA; 57516 MW; D16DD33BCCSF8EDA CRC64;
 Query Match 96.0%; Score 24; DB 1; Length 497;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 DXNDN 5
 Db 461 DSNDN 465
 RESULT 12
 LAC1_PLEOS STANDARD; PRT; 529 AA.
 ID_LAC1_PLEOS
 AC Q12729;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Laccase 1 precursor (EC 1.10.3.2) (Benzene-1,2-diol: oxygen oxidoreductase)
 DE (Urshiol oxidase) (Diphenol oxidase).
 GN POX1.
 OS Pleurotus ostreatus (Oyster mushroom) (White-rot fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Pleurotaceae; Pleurotus.
 OX NCBI_TaxID=5322;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Florida; TISSUE=Mycelium;
 RX MEDLINE=95314294; PubMed=7793361;
 RA Giardina P., Cannio R., Martirani L., Marzullo L., Palmieri G.,
 RA Sanna G.;
 RT "Cloning and sequencing of a laccase gene from the lignin-degrading
 RT basidiomycete Pleurotus ostreatus.";
 RL Appl. Environ. Microbiol. 61:2408-2413(1995).

CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
 CC PRODUCTS (PROBABLE).
 CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
 CC H(2)O.
 CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
 CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
 CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z34847; CAAB84356.1; -;
 CC EMBL; Z22591; CAAB80305.1; -;
 CC HSSP; P37064; LAOZ.
 CC InterPro; IPR001117; Cu-oxidase.
 CC InterPro; IPR002355; MultiCu oxidase2.
 CC Pfam; PF00394; Cu-oxidase; 3.
 CC PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
 CC PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 KW Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation;
 KW Glycoprotein; Repeat; Multigene family.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 529 LACCASE 1.
 FT DOMAIN 25 159 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 170 312 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 380 499 PLASTOCYANIN-LIKE 3.
 FT METAL 96 96 COPPER (TYPE 2) (BY SIMILARITY).
 FT METAL 98 98 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 141 141 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 143 143 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 425 425 COPPER (TYPE 1) (BY SIMILARITY).
 FT METAL 428 428 COPPER (TYPE 2) (BY SIMILARITY).
 FT METAL 430 430 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 481 481 COPPER (TYPE 1) (BY SIMILARITY).
 FT METAL 482 482 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 483 483 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 487 487 COPPER (TYPE 1) (BY SIMILARITY).
 FT METAL ? ? COPPER (TYPE 1) (BY SIMILARITY).
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 529 AA; 56580 MW; 40110AFC486882F CRC64;
 Query Match 96.0%; Score 24; DB 1; Length 529;
 Best Local Similarity 80.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 DXNDN 5
 Db 461 DANDN 465
 RESULT 13
 YGR0_YEAST STANDARD; PRT; 570 AA.
 ID_YGR0_YEAST
 AC PS3109;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 65.8 kDa protein in SUT1-RCK1 intergenic region.
 GN YGL160W OR G1837.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.


```

OX NCBI_TaxID=4932;
RN [1]
RX MEDLINE=96158061; PubMed=8585324;
RA James C.M., Inge K.J., Oliver S.G.;
RT "DNA sequence analysis of a 35 kb segment from Saccharomyces
RT cerevisiae chromosome VII reveals 19 open reading frames including
RT RAD54, ACE1/CUP2, PMR1, RCK1, AMS1 and CALL/CDC43."
RL Yeast 11:1413-1419(1995).
CC
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 272682; CAA96872.1; -
CC SGD; S0003128; YGL160W.
CC InterPro: IPR002916; Ferric_reduct.
CC Pfam; PF01794; Ferric_reduct. 1.
CC Hypothetical protein; Transmembrane.
FT TRANSMEM 21 41. POTENTIAL.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 142 162 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
SQ SEQUENCE 570 AA; 65839 MW; D2534C140A04FB8 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 570;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DNDN 5
DB 489 DSDN 493

RESULT 14
SP10 YEAST
ID SP10 YEAST STANDARD; PRT; 640 AA.
AC P35208;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SPT10 protein.
GN SPT10 OR SUD1 OR CRE1 OR YJL127C OR J0702.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A. AND MUTAGENESIS.
RX MEDLINE=94186069; PubMed=8138180;
RA Natsoulis G., Winston P., Boeke J.D.;
RT "The SP10 and SPT20 genes of Saccharomyces cerevisiae."
RL Genetics 136:93-105(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Cziepluch C., Kordes E., Fajol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon."
RL Yeast 12:1471-1474(1996).
RN [3]
RN SEQUENCE OF 80-640 FROM N.A.
RC STRAIN=GRF88;
RX MEDLINE=94088455; PubMed=8264536;

```

```

RA Yamashita I.;
RT "Isolation and characterization of the SUD1 gene, which encodes a
RT global repressor of core promoter activity in Saccharomyces
RT cerevisiae."
RL Mol. Gen. Genet. 241:616-626(1993)
CC -|- FUNCTION: REQUIRED FOR NORMAL TRANSCRIPTION AT A NUMBER OF LOCI IN
CC YEAST. AFFECTS TRANSCRIPTION AT TY1 ELEMENTS, AT PHOS, STE6 AND
CC ADH2.
CC -|- SIMILARITY: TO S.POMBE SPAC21E11.04.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L24435; AAA35077.1; -
CC EMBL; Z49402; CAA89422.1; -
CC EMBL; X65186; CAA46300.1; -
CC PIR; S47865; S47865.
CC TRANSFAC; T04377; -
CC SGD; S0003663; SPT10.
CC InterPro: IPR00182; GCN5acetyltransf.
CC Pfam; PF00583; Acetyltransf. 1.
FT MUTAGEN 388 C->S: LOSS OF COMPLEMENTATION OF THE
FT SUPPRESSION PHENOTYPE.
SQ SEQUENCE 640 AA; 72922 MW; 9E81C87E2A6D6C5 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 640;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DNDN 5
DB 130 DTNDN 134

RESULT 15
YFE2 YEAST
ID YFE2 YEAST STANDARD; PRT; 674 AA.
AC P43560; P43559;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 76.3 kDa protein in SEC53-FET5 intergenic region.
GN YFLO42C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RL Nat. Genet. 10:261-268(1995).
RN [2]
RN REVISIONS.
RP Murakami Y.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC -|- SIMILARITY: TO YEAST YHR080C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D50617; BAA24424.1; --
 DR SGD; S0001852; YFL042C.
 DR InterPro; IPR004182; GRAM_dom.
 DR Pfam; PF02893; GRAM; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 634 654 POTENTIAL.
 SQ SEQUENCE 674 AA; 76347 MW; 161F12303E1102C4 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 674;
 Best Local Similarity 80.0%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 DB 344 DANDN 348

RESULT 16
 YKF4_YEAST STANDARD; PRT; 738 AA.
 ID YKF4_YEAST
 AC P35732;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Hypothetical 84.0 kbp protein in NUP120-CS84 intergenic region.
 GN YKL054C OR YKL308.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=94378723; PubMed=8091862;
 RA Rasmussen S.W.;
 RT "Sequence of a 28.6 kb region of yeast chromosome XI includes the
 FT FBA1 and TOA2 genes, an open reading frame (ORF) similar to a
 RT translationally controlled tumour protein, one ORF containing motifs
 RT also found in plant storage proteins and 13 ORFs with weak or no
 RT homology to known proteins.";
 RL Yeast 10:563-568(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X75781; CAA53418.1; --
 DR EMBL; Z28054; CAA81890.1; --
 DR PIR; S37876; S37876.
 DR PIR; S39184; S39184.
 DR SGD; S0001537; YKL054C.
 KW Hypothetical protein.
 FT DOMAIN 382 718 GLN-RICH.
 FT DOMAIN 574 580 POLY-ALA.
 FT DOMAIN 708 716 POLY-ALA.
 SQ SEQUENCE 738 AA; 83973 MW; 05734E2D0B7389AC CRC64;

Query Match 96.0%; Score 24; DB 1; Length 738;
 Best Local Similarity 80.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 DB 177 DSNDN 181

RESULT 17
 KM65_YEAST STANDARD; PRT; 742 AA.
 ID KM65_YEAST
 AC Q03656;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable serine/threonine-protein kinase YMR216C (EC 2.7.1.-).
 GN YMR216C OR YMR261.10C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,
 Walsh S.V.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SIMILAR TO S.POMBE DSK1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; Z49809; CAA89931.1; --
 DR SGD; S0004829; YMR216C.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 158 706 PROTEIN KINASE.
 FT NP_BIND 164 172 ATP (BY SIMILARITY).
 FT BINDING 187 187 ATP (BY SIMILARITY).
 FT ACT_SITE 294 294 BY SIMILARITY.
 SQ SEQUENCE 742 AA; 83237 MW; C775F10B30C950FC CRC64;

Query Match 96.0%; Score 24; DB 1; Length 742;
 Best Local Similarity 80.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 DB 462 DSNDN 466,

RESULT 18
 YLDM_CAEEL STANDARD; PRT; 758 AA.
 ID YLDM_CAEEL
 AC Q94225;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein F41D9.5 in chromosome X.
 GN F41D9.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Fulton B.;

Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
[2]
REVISIONS.
RA Waterston R.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: POSSIBLE SULFATE TRANSPORTER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.
CC -!- SIMILARITY: CONTAINS 1 STAS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67954; AAB52608.1; -;
DR WormPep; F41D9.5; CE10242.
DR InterPro; IPR002645; STAS.
DR Pfam; PF00916; Sulfate_transp; 1.
DR Pfam; PF01740; STAS; 1.
DR PROSITE; PS01130; SLC26A; 1.
DR PROSITE; PS0801; STAS; 1.
KW Hypothetical protein; Transmembrane; Transport.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 60 80 POTENTIAL.
FT TRANSMEM 82 102 POTENTIAL.
FT TRANSMEM 109 129 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 228 248 POTENTIAL.
FT TRANSMEM 270 290 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT TRANSMEM 335 355 POTENTIAL.
FT TRANSMEM 375 395 POTENTIAL.
FT TRANSMEM 409 429 POTENTIAL.
FT TRANSMEM 473 493 POTENTIAL.
FT DOMAIN 522 676 STAS.
SQ SEQUENCE 758 AA; 84958 MW; 721B87729F2E12F3 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 758;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 710 DSDND 714

RESULT 19
CDBG HUMAN STANDARD; PRT; 776 AA.
AC Q9NRJ7; O9HCF1.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin beta 16 precursor (PCDH-beta16) (Protocadherin 3X).
GN PCDHB16 OR PCDH3X OR KIA1621.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kools P.F.J., Van Roy F.;
RT "The human protocadherin 3 gene cluster, expression analysis of these
RT one-exon genes."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;

MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hiroawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Kools P.F.J., van Roy F.M.;
RT "Molecular analysis of the human protocadherin-3 (PCDH-Beta) gene
RT cluster."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF217757; AAF81914.1; -;
DR EMBL; AB046841; BAB13447.1; ALT_INIT.
DR EMBL; AF282973; AAG10030.1; -;
DR Genew; HGNC:14546; PCDHB16.
DR MIM; 606345; -;
DR MIM; 604967; -;
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN 1; 5.
DR PROSITE; PS0268; CADHERIN 2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;
KW Transmembrane; Multigene family.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 776 PROTOCADHERIN BETA 16.
FT DOMAIN 29 690 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 712 776 POTENTIAL.
FT DOMAIN 712 776 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 35 133 CADHERIN 1.
FT DOMAIN 138 242 CADHERIN 2.
FT DOMAIN 247 347 CADHERIN 3.
FT DOMAIN 352 451 CADHERIN 4.
FT DOMAIN 456 551 CADHERIN 5.
FT DOMAIN 568 671 CADHERIN 6.
FT CARBOHYD 418 418 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 482 482 T -> I (IN REF. 2).
FT CONFLICT 525 526 QA -> RE (IN REF. 2).
FT CONFLICT 532 532 G -> S (IN REF. 2).
FT CONFLICT 543 543 R -> S (IN REF. 2).
FT CONFLICT 638 638 Q -> H (IN REF. 2).
FT CONFLICT 652 652 R -> C (IN REF. 2).
FT CONFLICT 710 710 A -> V (IN REF. 2).
SQ SEQUENCE 776 AA; 84935 MW; 93BF1C641A1DE0FB CRC64;

Query Match 96.0%; Score 24; DB 1; Length 776;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 553 DANDN 557

RESULT 20
CDBF HUMAN STANDARD; PRT; 787 AA.
ID CDBF_HUMAN

AC 09Y588;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protocadherin beta 15 precursor (PCDH-beta15).
 GN PCDHB15.
 OS Homo sapiens (HUMAN).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99308636; PubMed=10380929;
 RA Wu Q., Maniatis T.;
 RT "A striking organization of a large family of human neural cadherin-
 RL like cell adhesion genes.";
 RN Cell 97:779-790 (1999).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21223055; PubMed=11222959;
 RA Vanhalst K., Kools P., Vanden Eynde E., van Roy F.;
 RT "The human and murine protocadherin-beta one-exon gene families show
 RT high evolutionary conservation, despite the difference in gene
 RL number.";
 RN FEBS Lett. 495:120-125 (2001).
 CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF152494; AAD43755.1; -;
 CC EMBL; AF217742; AAKS1610.1; -;
 CC Genew; HGNC:8686; PCDHB15.
 CC MIM; 606341; -;
 CC MIM; 604967; -;
 CC InterPro; IPR002126; Cadherin.
 CC Pfam; PF00028; cadherin; 5.
 CC PRINTS; PR00205; CADHERIN.
 CC SMART; SM00112; CA; 5.
 CC PROSITE; PS00232; CADHERIN 1; 5.
 CC PROSITE; PS0268; CADHERIN 2; 6.
 CC Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;
 KW Transmembrane, Multigene family.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 787 PROTOCADHERIN BETA 15.
 FT DOMAIN 27 690 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 691 711 POTENTIAL.
 FT DOMAIN 712 787 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 35 133 CADHERIN 1.
 FT DOMAIN 138 242 CADHERIN 2.
 FT DOMAIN 247 347 CADHERIN 3.
 FT DOMAIN 352 451 CADHERIN 4.
 FT DOMAIN 456 561 CADHERIN 5.
 FT DOMAIN 568 671 CADHERIN 6.
 FT CARBOHYD 418 418 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 567 567 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 787 AA; 86328 MW; 8D8D3A07B086A56 CRC64;
 Query Match 96.0%; Score 24; DB 1; Length 787;
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DNDN 5
 DB 234 DANDN 238

RESULT 21
 CADA HUMAN STANDARD; PRT; 788 AA.
 AC 09Y6N8; OSULB3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-10 precursor (T2-cadherin).
 GN CDH10.
 OS Homo sapiens (HUMAN).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=99313187; PubMed=10386616;
 RA Kools P., Vanhalst K., Van den Eynde E., Van Roy F.;
 RT "The human cadherin-10 gene: complete coding sequence, predominant
 RT expression in the brain, and mapping on chromosome 5p13-14.";
 RN FEBS Lett. 452:328-334 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=10861224;
 RA Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;
 RT "Identification of three human type-II classic cadherins and frequent
 RT heterophilic interactions between different subclasses of type-II
 RT classic cadherins.";
 RL Biochem. J. 349:159-167 (2000).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN. ALSO FOUND
 CC IN ADULT AND FETAL KIDNEY. VERY LOW LEVELS DETECTED IN PROSTATE
 CC AND FETAL LUNG.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF039747; AAD44017.1; -;
 CC EMBL; AB035303; BAA87417.1; -;
 CC HSPSP; F15116; INCJ.
 CC Genew; HGNC:1749; CDH10.
 CC MIM; 604555; -;
 CC InterPro; IPR002126; Cadherin.
 CC InterPro; IPR000233; Cadherin_C-term.
 CC Pfam; PF00028; cadherin; 5.
 CC Pfam; PF01049; Cadherin_C-term; 1.
 CC PRINTS; PR00205; CADHERIN.
 CC SMART; SM00112; CA; 5.
 CC PROSITE; PS00232; CADHERIN 1; 3.
 CC PROSITE; PS0268; CADHERIN 2; 5.
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 54 POTENTIAL.
 FT CHAIN 55 788 CADHERIN-10.
 FT DOMAIN 55 613 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 614 634 POTENTIAL.
 FT DOMAIN 635 788 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 55 160 CADHERIN 1.
 FT DOMAIN 161 269 CADHERIN 2.
 FT DOMAIN 270 384 CADHERIN 3.
 FT DOMAIN 385 487 CADHERIN 4.
 FT DOMAIN 488 606 CADHERIN 5.

```
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 34 34 R -> Q (IN REF. 2).
FT CONFLICT 482 482 A -> V (IN REF. 2).
SQ SEQUENCE 788 AA; 88450 MW; C6F116E2CF0D68CA CRC64;

Query Match 96.0%; Score 24; DB 1; Length 788;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 481 DANDN 485

RESULT 22
CDB7 HUMAN STANDARD; PRT; 793 AA.
ID CDB7 HUMAN STANDARD; PRT; 793 AA.
AC QY5E2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin beta 7 precursor (PCDH-beta7).
GN PCDH87.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=99308636; PubMed=10380929;
RA Wu Q., Maniatis T.;
RT "A striking organization of a large family of human neural cadherin-
RT like cell adhesion genes.";
RL Cell 97:779-790(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223055; PubMed=11322959;
RA Vanhalst K., Kools P., Vanden Eynde E., van Roy F.;
RT "The human and murine protocadherin-beta one-exon gene families show
RT high evolutionary conservation, despite the difference in gene
RT number.";
RL FEBS Lett. 495:120-125(2001).
CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF152500; AAD43761.1; -
CC EMBL; AF217750; AAK51617.1; -
CC HSP; P15116; INCU.
CC Genew; HGNC:8692; PCDH87.
CC MIM; 606333; -
CC MIM; 604967; -
CC InterPro; IPR002126; Cadherin.
CC Pfam; PF00028; cadherin; 5.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 5.
CC PROSITE; PS00232; CADHERIN_1; 5.
CC PROSITE; PS00268; CADHERIN_2; 5.
CC Calcium-binding; Call adhesion; Glycoprotein; Signal; Repeat;
CC Transmembrane; Multigene family.
CC SIGNAL 26
CC CHAIN 27 793 PROTOCADHERIN BETA 7.
FT
```

```
FT DOMAIN 27 688 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 689 709 POTENTIAL.
FT DOMAIN 710 793 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 35 133 CADHERIN 1.
FT DOMAIN 138 242 CADHERIN 2.
FT DOMAIN 247 347 CADHERIN 3.
FT DOMAIN 352 451 CADHERIN 4.
FT DOMAIN 456 561 CADHERIN 5.
FT DOMAIN 568 671 CADHERIN 6.
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 793 AA; 86706 MW; E5677098653CC755 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 793;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 553 DANDN 557

RESULT 23
CDB6 HUMAN STANDARD; PRT; 794 AA.
ID CDB6 HUMAN STANDARD; PRT; 794 AA.
AC QY1SE3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin beta 6 precursor (PCDH-beta6).
GN PCDH86.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=99308636; PubMed=10380929;
RA Wu Q., Maniatis T.;
RT "A striking organization of a large family of human neural cadherin-
RT like cell adhesion genes.";
RL Cell 97:779-790(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223055; PubMed=11322959;
RA Vanhalst K., Kools P., Vanden Eynde E., van Roy F.;
RT "The human and murine protocadherin-beta one-exon gene families show
RT high evolutionary conservation, despite the difference in gene
RT number.";
RL FEBS Lett. 495:120-125(2001).
CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF152499; AAD43760.1; -
CC EMBL; AF217752; AAK51619.1; -
CC HSP; P15116; INCU.
CC Genew; HGNC:8691; PCDH86.
CC MIM; 606332; -
CC MIM; 604967; -
CC InterPro; IPR002126; Cadherin.
CC Pfam; PF00028; cadherin; 5.
CC PRINTS; PR00205; CADHERIN.
```



```

RC TISSUE=Uterus;
RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF152498; AD43759.1; -
DR EMBL; AF217753; AK51620.1; -
DR EMBL; BC001186; AH01186.1; -
DR EMBL; AL117449; CAB55933.1; -
DR Genew; HGNC:8690; PCDH5.
DR MIM; 606331; -
DR MIM; 604967; -
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN 1; 5.
DR PROSITE; PS0268; CADHERIN 2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;
KW Transmembrane; Multigene family.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 795 PROTOCADHERIN BETA 5.
FT DOMAIN 31 689 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 690 710 POTENTIAL.
FT DOMAIN 711 795 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 138 242 CADHERIN 1.
FT DOMAIN 247 346 CADHERIN 2.
FT DOMAIN 351 450 CADHERIN 3.
FT DOMAIN 455 560 CADHERIN 4.
FT DOMAIN 567 670 CADHERIN 5.
FT DOMAIN 671 795 CADHERIN 6.
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 566 566 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 720 720 S -> P (IN REF. 3 AND 4).
SQ SEQUENCE 795 AA; 86422 MW; 06F743580DC65B64 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 795;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 552 DANDN 556

Search completed: June 20, 2003, 20:52:49
Job time : 11.0227 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:42:49 ; Search time 26.3636 Seconds
(without alignments)
39.078 Million cell updates/sec

Title: US-10-105-008-8

Perfect score: 25

Sequence: 1 DKNND 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database :

SPREMBL 21:.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertibrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	96.0	89	12 Q84490	Q84490 paramecium
2	24	96.0	101	15 Q71280	Q71280 human immun
3	24	96.0	103	12 Q8QY8	Q8QY8 feline cali
4	24	96.0	105	5 Q22850	Q22850 caenorhabdi
5	24	96.0	108	6 Q8WJ36	Q8WJ36 macaca mula
6	24	96.0	110	10 Q94H27	Q94H27 oryza sativ
7	24	96.0	130	5 Q20980	Q20980 caenorhabdi
8	24	96.0	138	6 Q8WMJ0	Q8WMJ0 macaca mula
9	24	96.0	139	16 Q92JJ0	Q92JJ0 rickettsia
10	24	96.0	144	16 Q8XJ00	Q8XJ00 clostridium
11	24	96.0	146	6 Q8WNIS	Q8WNIS macaca mula
12	24	96.0	152	16 Q99Q29	Q99Q29 staphylococ
13	24	96.0	153	10 Q43491	Q43491 hordeum vul
14	24	96.0	166	16 Q8XVP0	Q8XVP0 ralestonia s
15	24	96.0	168	6 Q8WMH9	Q8WMH9 macaca mula
16	24	96.0	179	5 Q95P73	Q95P73 haematobia

ALIGNMENTS

RESULT 1
ID Q84490 PRELIMINARY; PRT; 89 AA.
AC Q84490;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE A170L protein.
GN A170L.

17	24	96.0	193	2	P70820	P70820 borrelia bu
18	24	96.0	195	15	Q8QAA3	Q8QAA3 human immun
19	24	96.0	195	15	Q8QAB6	Q8QAB6 human immun
20	24	96.0	196	15	Q8QAB6	Q8QAB6 human immun
21	24	96.0	196	15	Q8QAD9	Q8QAD9 human immun
22	24	96.0	196	15	Q8QAD7	Q8QAD7 human immun
23	24	96.0	196	15	Q8QAC1	Q8QAC1 human immun
24	24	96.0	196	15	Q8QAB5	Q8QAB5 human immun
25	24	96.0	197	15	Q8QAG4	Q8QAG4 human immun
26	24	96.0	197	15	Q8QAB3	Q8QAB3 human immun
27	24	96.0	199	15	Q8QAP5	Q8QAP5 human immun
28	24	96.0	199	15	Q8QAB8	Q8QAB8 human immun
29	24	96.0	199	15	Q8QAB0	Q8QAB0 human immun
30	24	96.0	199	15	Q8QAB8	Q8QAB8 human immun
31	24	96.0	201	10	Q8LYA4	Q8LYA4 arabidopsis
32	24	96.0	205	15	Q8US53	Q8US53 human immun
33	24	96.0	210	13	Q8QFV9	Q8QFV9 pagochenia
34	24	96.0	223	3	Q9HF76	Q9HF76 ashbya goes
35	24	96.0	225	2	Q9R3L9	Q9R3L9 actinobacil
36	24	96.0	226	2	Q9R764	Q9R764 actinobacil
37	24	96.0	238	5	Q9SP77	Q9SP77 halocynthia
38	24	96.0	243	10	Q9C7H1	Q9C7H1 arabidopsis
39	24	96.0	260	4	Q9H349	Q9H349 homo sapien
40	24	96.0	267	3	Q01138	Q01138 mucor circi
41	24	96.0	271	5	Q9N8A7	Q9N8A7 trypanosoma
42	24	96.0	273	10	Q9FEI4	Q9FEI4 avena byzan
43	24	96.0	280	16	Q97ES2	Q97ES2 clostridium
44	24	96.0	282	2	Q51156	Q51156 neisseria m
45	24	96.0	290	12	Q91EX4	Q91EX4 cydia pomon
46	24	96.0	295	3	Q12515	Q12515 saccharomyc
47	24	96.0	296	5	Q9NF63	Q9NF63 caenorhabdi
48	24	96.0	313	16	P94522	P94522 bacillus su
49	24	96.0	314	4	Q96RE3	Q96RE3 homo sapien
50	24	96.0	314	5	Q9VHM6	Q9VHM6 drosophila
51	24	96.0	315	2	Q9L5V0	Q9L5V0 salmonella
52	24	96.0	317	17	Q30168	Q30168 archaeoglob
53	24	96.0	323	16	Q9RGU8	Q9RGU8 salmonella
54	24	96.0	324	2	Q07078	Q07078 bacillus su
55	24	96.0	328	5	Q9VDU6	Q9VDU6 drosophila
56	24	96.0	328	13	Q42244	Q42244 xenopus lae
57	24	96.0	330	2	P77860	P77860 erwinia car
58	24	96.0	351	16	Q9PG86	Q9PG86 xylella fas
59	24	96.0	354	5	Q8SSW8	Q8SSW8 dictyosteli
60	24	96.0	357	5	Q76412	Q76412 caenorhabdi
61	24	96.0	360	3	Q96VA2	Q96VA2 metarhizium
62	24	96.0	366	2	Q44163	Q44163 actinobacil
63	24	96.0	367	2	Q9ZAP9	Q9ZAP9 actinobacil
64	24	96.0	367	2	Q44164	Q44164 actinobacil
65	24	96.0	382	4	Q96P92	Q96P92 homo sapien
66	24	96.0	389	16	Q9KOT6	Q9KOT6 vibrio chol
67	24	96.0	390	3	Q96UG1	Q96UG1 metarhizium
68	24	96.0	390	4	Q96KB8	Q96KB8 homo sapien
69	24	96.0	399	16	Q92Q24	Q92Q24 rhizobium m
70	24	96.0	403	15	O11581	O11581 human immun
71	24	96.0	413	5	Q9VXE2	Q9VXE2 drosophila
72	24	96.0	420	4	O95883	O95883 homo sapien
73	24	96.0	422	17	Q96Y13	Q96Y13 sulfolobus
74	24	96.0	426	16	Q8VJL6	Q8VJL6 mycobacteri
75	24	96.0	440	16	P71894	P71894 mycobacteri

OS Paramesitium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95407089; PubMed=7676624;
RA Li Y., Lu Z., Burbank D.E., Kutish G.F., Rock D.L., Van Etten J.L.;
RT "Analysis of 43 kb of the Chlorella virus PBCV-1 330-Kb genome: map
RT positions 45 to 88."
RL Virology 212:134-150(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20033326; PubMed=10544099;
RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA Lisec A.D., Nickerson K.W., Van Etten J.L.;
RT "Chlorella virus PBCV-1 encodes a functional homospesmidine
RT synthase."
RL Virology 263:254-262(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20478054; PubMed=11021991;
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RT PBCV-1."
RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Gurnon J.R., Graves M.V., Van Etten J.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U42580; AAC96538.1; -.
SQ SEQUENCE 89 AA; 9855 MW; 2006F23083AADFFE CRC64;
Query Match 96.0%; Score 24; DB 12; Length 89;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DXNDN 5
Db 43 DSNDN 47
RESULT 2
ID Q71280 PRELIMINARY; PRT; 101 AA.
AC Q71280;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
OS Paramesitium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CASE B;
RA Wang N., Zhu T., Ho D.D.;
RT "Sequence diversity of V1 and V2 domains of gp120 from human
RT immunodeficiency virus type 1: lack of correlation with viral
RT phenotype."
RL J. Virol. 69:2708-2715(1995).
DR EMBL; U19650; AAA68667.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 101
SQ SEQUENCE 101 AA; 11280 MW; BCB994FF648B085B CRC64;
Query Match 96.0%; Score 24; DB 15; Length 101;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DXNDN 5
Db 62 DANDN 66
RESULT 3
ID Q8QY8 PRELIMINARY; PRT; 103 AA.
AC Q8QY8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Capsid (Fragment).
OS Feline calicivirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11978;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S183;
RA Radford A.D., Dawson S., Ryvar R., Johnson D.R., Cox M.B., Acke E.,
RA Addie D.D., Gaskell R.M.;
RT "Evolution of feline calicivirus."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF366229; AAL99720.1; -.
FT NON_TER 1
FT NON_TER 103
SQ SEQUENCE 103 AA; 11422 MW; CF573A54A1EDE7F CRC64;
Query Match 96.0%; Score 24; DB 12; Length 103;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DXNDN 5
Db 20 DSNDN 24
RESULT 4
ID Q22850 PRELIMINARY; PRT; 105 AA.
AC Q22850;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE T28F4.5 protein.
GN T28F4.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; Z72517; CAA96692.1; -;
SQ SEQUENCE 105 AA; 11957 MW; 04636B3281A3BB2C CRC64;

Query Match 96.0%; Score 24; DB 5; Length 105;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 36 DSNDN 40

RESULT 5
Q8WMJ6 PRELIMINARY; PRT; 108 AA.
ID Q8WMJ6
AC Q8WMJ6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Protocadherin alpha 7 (Fragment).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MEDIAL BASAL HYPOTHALAMUS;
RA Brown A.E., Ojeda S.R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF424821; AAL40037.1; -;
DR InterPro; IPR002126; Cadherin.
DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_1.
DR PROSITE; PS50268; CADHERIN_2; 1.
FT NON_TER 108
SQ SEQUENCE 108 AA; 11529 MW; 86B202422E127F4E CRC64;

Query Match 96.0%; Score 24; DB 6; Length 108;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 48 DANDN 52

RESULT 6
Q94HZ7 PRELIMINARY; PRT; 110 AA.
ID Q94HZ7
AC Q94HZ7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 11.9 kDa protein.
OS Oryza sativa (rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,

RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,
RA Uterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNBa0076F20 genomic sequence."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025296; AAK39578.1; -;
KW Hypothetical protein.
SQ SEQUENCE 110 AA; 11947 MW; 4FD6590E4615888D CRC64;

Query Match 96.0%; Score 24; DB 10; Length 110;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 8 DSNDN 12

RESULT 7
Q20980 PRELIMINARY; PRT; 130 AA.
ID Q20980
AC Q20980;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F5886.4 protein.
GN F5886.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; Z70754; CAA94774.1; -;
SQ SEQUENCE 130 AA; 14628 MW; 5ADDA9087D54F0F0 CRC64;

Query Match 96.0%; Score 24; DB 5; Length 130;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 94 DTNDN 98

RESULT 8
Q8WMJ0 PRELIMINARY; PRT; 138 AA.
ID Q8WMJ0
AC Q8WMJ0;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Protocadherin alpha C1 (Fragment).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MEDIAL BASAL HYPOTHALAMUS;
RA Brown A.E., Ojeda S.R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF424828; AAL40243.1; -;
 DR InterPro; IPR002126; Cadherin.
 DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_1.
 DR PROSITE; PS0268; CADHERIN_2; 1.
 FT NON_TER 138 138
 FT NON_TER 138 138
 SQ SEQUENCE 138 AA; 14998 MW; AB3BB21FFE35EB48 CRC64;

Query Match 96.0%; Score 24; DB 6; Length 138;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 113 DTNDN 117

RESULT 9

ID Q92JJO PRELIMINARY; PRT; 139 AA.
 AC Q92JJO
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Hypothetical protein RCO077.
 GN RCO077.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MALISH 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Saeon D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
 RL Science 293:2093-2098(2001).
 DR EMBL; AE008575; AAL02615.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 139 AA; 15918 MW; 036F69A5C85CBEEA CRC64;

Query Match 96.0%; Score 24; DB 16; Length 139;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 129 DSNDN 133

RESULT 10

ID Q8XJ00 PRELIMINARY; PRT; 144 AA.
 AC Q8XJ00
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Hypothetical protein CPE1963.
 GN CPE1963.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 RX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic flesh-eater";

RL PROC. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL; AP003192; BAB81669.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 144 AA; 15678 MW; BB770D050ED551DB CRC64;

Query Match 96.0%; Score 24; DB 16; Length 144;
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 47 DTNDN 51

RESULT 11

ID Q8WMIS PRELIMINARY; PRT; 146 AA.
 AC Q8WMIS
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Protocadherin beta 3' (Fragment).
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MEDIAL BASAL HYPOTHALAMUS;
 RA Brown A.E., Ojeda S.R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF424834; AAL40248.1; -;
 DR InterPro; IPR002126; Cadherin.
 PFam; PF00028; cadherin; 1.
 DR SMART; SM00112; CA; 1.
 DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_1.
 DR PROSITE; PS0268; CADHERIN_2; 1.
 FT NON_TER 146 146
 FT NON_TER 146 146
 SQ SEQUENCE 146 AA; 15501 MW; D2F8B904FA845099 CRC64;

Query Match 96.0%; Score 24; DB 6; Length 146;
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 105 DANDN 109

RESULT 12

ID Q99QZ9 PRELIMINARY; PRT; 152 AA.
 AC Q99QZ9
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE Hypothetical protein SAV2639.
 GN SAV2639 OR SA2432.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=158878; 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
 RT "Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*."
 RL Lancet 357:1225-1240 (2001).
 DR EMBL; AF003366; BAB58801.1; -.
 DR EMBL; AF003138; BAB43737.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 152 AA; 16363 MW; 2C79EC18ACBB9PBF CRC64;

Query Match 96.0%; Score 24; DB 16; Length 152;
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 DB 71 DTNDN 75

RESULT 13
 Q43491 PRELIMINARY; PRT; 153 AA.
 ID Q43491
 AC Q43491
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 17.7 kDa protein.
 OS Hordeum vulgare (barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxID=4513;
 RN [1]_TaxID=4513;
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CI-16137; TISSUE=LEAF;
 RX MEDLINE=97137848; PubMed=8983185;
 RA Liu K., Somerville S.;
 RT "Cloning and characterization of a highly repeated DNA sequence in
 RT Hordeum vulgare L.";
 RL Genome 39:1159-1168 (1996).
 DR EMBL; L23524; AAB42154.1; -.
 SQ SEQUENCE 153 AA; 17706 MW; 096F57A60E95E428 CRC64;

Query Match 96.0%; Score 24; DB 10; Length 153;
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 DB 66 DANDN 70

RESULT 14
 Q8XVP0 PRELIMINARY; PRT; 166 AA.
 ID Q8XVP0
 AC Q8XVP0
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Probable thiol peroxidase-scavengase p20 oxidoreductase protein
 DE (EC 1.11.1.-).
 GN TPX OR RSC2790 OR RS00057.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OC NCBI_TaxID=305;
 RN [1]_TaxID=305;
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,

RA Chandler M., Choiane N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Signier P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*."
 RL Nature 415:497-502 (2002).
 DR EMBL; AL646071; CAD16497.1; -.
 DR InterPro; IPR000866; AhpC-TSA.
 DR InterPro; IPR002065; TPX.
 DR Pfam; PF00578; AhpC-TSA; 1.
 DR PROSITE; PS01265; TPX; 1.
 KW Oxidoreductase; Peroxidase; Complete proteome.
 SQ SEQUENCE 166 AA; 17283 MW; 6621B12E5F391F13 CRC64;

Query Match 96.0%; Score 24; DB 16; Length 166;
 Best Local Similarity 80.0%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 DB 138 DANDN 142

RESULT 15
 Q8WMH9 PRELIMINARY; PRT; 168 AA.
 ID Q8WMH9
 AC Q8WMH9
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Protocadherin alpha 5 (fragment).
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBI_TaxID=9544;
 RN [1]_TaxID=9544;
 RP SEQUENCE FROM N.A.
 RC TISSUE=MEDIAL BASAL HYPOTHALAMUS;
 RA Mungenast A.E., Oyeda S.R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF424840; AAL40254.1; -.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 1.
 DR SMART; SM00112; CA; 1.
 DR PROSITE; PS00232; CADHERIN_1; UNKNOWN 2.
 DR PROSITE; PS0268; CADHERIN_2; 2.
 DR NON TER 1
 FT NON TER 168
 SQ SEQUENCE 168 AA; 19305 MW; 849884540B073E8E CRC64; 1*

Query Match 96.0%; Score 24; DB 6; Length 168;
 Best Local Similarity 80.0%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
 DB 161 DANDN 165

RESULT 16
 Q95P73 PRELIMINARY; PRT; 179 AA.
 ID Q95P73
 AC Q95P73
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Thrombostasin (fragment).
 OS Haematobia irritans (horn fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Muscoidea; Muscidae; Haematobia.
 OC NCBI_TaxID=7368;

```

RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,
RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
RA Tobin N.H., Mullins J.I.;
RT "Potency of antiretroviral therapy affects viral evolution and the
RT efficacy of HIV-1 therapy.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY076822; AAL71223.1; -.
FT NON_TER 1
FT NON_TER 195
SQ SEQUENCE 195 AA; 22023 MW; F86352AC6613F033 CRC64;

Query Match 96.0%; Score 24; DB 5; Length 179;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 128 DSNDN 132

RESULT 17
ID P70820 PRELIMINARY; PRT; 193 AA.
AC P70820;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Surface-exposed lipoprotein (fragment).
GN NLPH.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DK7;
RC MEDLINE=97086609; PubMed=8932298;
RA "Thaisen M.; cloning and characterization of nlph, encoding a novel
RT surface-exposed, polymorphic, 33-kilodalton lipoprotein of Borrelia
RT afzelii.";
RL J. Bacteriol. 178:6435-6435(1996).
DR EMBL; Y08412; CAA69688.1; -.
DR InterPro; IPR004983; Mlp.
DR Pfam; PF03304; Mlp; 1.
KW Lipoprotein.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 22567 MW; 99D2454A4322C44D CRC64;

Query Match 96.0%; Score 24; DB 2; Length 193;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 2 DSNDN 6

RESULT 18
ID Q8QAA3 PRELIMINARY; PRT; 195 AA.
AC Q8QAA3;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Envelope glycoprotein (fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99USL999616ECB;

RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,
RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
RA Tobin N.H., Mullins J.I.;
RT "Potency of antiretroviral therapy affects viral evolution and the
RT efficacy of HIV-1 therapy.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY076822; AAL71223.1; -.
FT NON_TER 1
FT NON_TER 195
SQ SEQUENCE 195 AA; 22023 MW; F86352AC6613F033 CRC64;

Query Match 96.0%; Score 24; DB 15; Length 195;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 129 DTNDN 133

RESULT 19
ID Q8QAA98 PRELIMINARY; PRT; 195 AA.
AC Q8QAA98;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Envelope glycoprotein (fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99USL999616ECG;
RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,
RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
RA Tobin N.H., Mullins J.I.;
RT "Potency of antiretroviral therapy affects viral evolution and the
RT efficacy of HIV-1 therapy.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY076827; AAL71228.1; -.
FT NON_TER 1
FT NON_TER 195
SQ SEQUENCE 195 AA; 22039 MW; 02D788346DDA7125 CRC64;

Query Match 96.0%; Score 24; DB 15; Length 195;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DXNDN 5
Db 129 DTNDN 133

RESULT 20
ID Q8QAE6 PRELIMINARY; PRT; 196 AA.
AC Q8QAE6;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Envelope glycoprotein (fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=92USL192916ECF;
RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
```

RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,
RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
RA Tobin N.H., Mullins J.I.,
RT "Potency of antiretroviral therapy affects viral evolution and the
RT efficacy of HIV-1 therapy.",
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY076779; AAL71180.1; -
FT NON_TER 1
FT SEQUENCE 196 AA; 22044 MW; 428D4EFC6AAD8B5D CRC64;

Query Match 96.0%; Score 24; DB 15; Length 196;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 131 DTNDN 135

RESULT 21

Q8QAD9 PRELIMINARY; PRT; 196 AA.
ID Q8QAD9
AC Q8QAD9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=92USL1929XEC16;
RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,
RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
RA Tobin N.H., Mullins J.I.,
RT "Potency of antiretroviral therapy affects viral evolution and the
RT efficacy of HIV-1 therapy.",
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY076786; AAL71187.1; -
FT NON_TER 1
FT SEQUENCE 196 AA; 22133 MW; BED37AD01013DB5A CRC64;

Query Match 96.0%; Score 24; DB 15; Length 196;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 131 DTNDN 135

RESULT 22

Q8QAD7 PRELIMINARY; PRT; 196 AA.
ID Q8QAD7
AC Q8QAD7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=92USL1929XEC18;
RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,

RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
RA Tobin N.H., Mullins J.I.,
RT "Potency of antiretroviral therapy affects viral evolution and the
RT efficacy of HIV-1 therapy.",
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY076788; AAL71189.1; -
FT NON_TER 1
FT SEQUENCE 196 AA; 22032 MW; FAB3485EE09959C5 CRC64;

Query Match 96.0%; Score 24; DB 15; Length 196;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 131 DTNDN 135

RESULT 23

Q8QAC1 PRELIMINARY; PRT; 196 AA.
ID Q8QAC1
AC Q8QAC1;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98USL198527ECD;
RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,
RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
RA Tobin N.H., Mullins J.I.,
RT "Potency of antiretroviral therapy affects viral evolution and the
RT efficacy of HIV-1 therapy.",
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY076804; AAL71205.1; -
FT NON_TER 1
FT SEQUENCE 196 AA; 21958 MW; 2CFD5E56B42307A5 CRC64;

Query Match 96.0%; Score 24; DB 15; Length 196;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DXNDN 5
DB 131 DTNDN 135

RESULT 24

Q8QAB5 PRELIMINARY; PRT; 196 AA.
ID Q8QAB5
AC Q8QAB5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98USL198527ECM;
RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
RA McKernan J.L., De Vange S.M., Holte S., Sylva G.L., Naugler W.E.,
RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,

RA Tobin N.H., Mullins J.I.;
 RT "Potency of antiretroviral therapy affects viral evolution and the
 efficacy of HIV-1 therapy."; EMBL/GenBank/DBJ databases.
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY076810; AAL71211.1; -
 FT NON_TER 1
 FT NON_TER 196
 SQ SEQUENCE 196 AA; 22035 MW; 47966C7CE88FEAA8 CRC64;

Query Match 96.0%; Score 24; DB 15; Length 196;
 Best Local Similarity 80.0%; Pred. NO. 4.9e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 131 DTNDN 135

RESULT 25

Q8QAG4 PRELIMINARY; PRT; 197 AA.
 AC Q8QAG4;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=01USL101214EC2;
 RA Frenkel L.M., Wang Y., Learn G.H., Melvin A.J., Ellis G.M.,
 RA McKernan J.L., De Vange S.M., Hoite S., Sylva G.L., Naugler W.E.,
 RA Mahalanabis M., Beck I., Mohan K.M., Lewis P.A., Heath L.M.,
 RA Tobin N.H., Mullins J.I.;
 RT "Potency of antiretroviral therapy affects viral evolution and the
 efficacy of HIV-1 therapy."; EMBL/GenBank/DBJ databases.
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY076761; AAL71162.1; -
 FT NON_TER 1
 FT NON_TER 197
 SQ SEQUENCE 197 AA; 22136 MW; F7C72901492169F8 CRC64;

Query Match 96.0%; Score 24; DB 15; Length 197;
 Best Local Similarity 80.0%; Pred. NO. 5e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXNDN 5
 Db 131 DTNDN 135

Search completed: June 20, 2003, 20:56:48
 Job time : 30.3636 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 38.3182 Seconds
(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-10

Perfect score: 41

Sequence: 1 YCHAVC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : A_Geneseq_101002.*

```
1: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	41	100.0	6 23	AAW47521
2	41	100.0	61 21	AA198085
3	41	100.0	61 21	AA198110
4	38	92.7	61 22	AAO09658
5	36	87.8	187 22	AA198984
6	36	87.8	428 23	AB192762
7	36	87.8	519 21	AA192468
8	36	87.8	519 22	AA192468
9	35	85.4	352 21	AA192468
10	34	82.9	5 19	AAW23979

RESULT 1
AAW47521
ID AAW47521 standard; peptide; 6 AA.

ALIGNMENTS

11	34	82.9	5	20	AAW17108	Cadherin-mediated
12	34	82.9	5	20	AAW91021	Peptide inhibiting
13	34	82.9	5	20	AAW91033	Peptide inhibiting
14	34	82.9	5	21	AAW27334	Beta-catenin deriv
15	34	82.9	5	21	AAW27346	Beta-catenin deriv
16	34	82.9	5	21	AAW73790	N-cadherin cell ad
17	34	82.9	5	22	AAW65370	Cyclic peptide wit
18	34	82.9	5	23	AAW47477	Cyclic peptide end
19	34	82.9	5	23	AAW65441	Cyclic peptide wit
20	34	82.9	6	22	AAW65444	Cyclic peptide wit
21	34	82.9	6	22	AAW65445	Cyclic peptide wit
22	34	82.9	6	22	AAW65448	Cyclic peptide wit
23	34	82.9	6	22	AAW65449	Cyclic peptide wit
24	34	82.9	6	22	AAW65450	Cyclic peptide wit
25	34	82.9	6	23	AAW47551	Peptide, SEQ ID 84
26	34	82.9	6	23	AAW47554	Cyclic peptide end
27	34	82.9	6	23	AAW47555	Cyclic peptide end
28	34	82.9	6	23	AAW47558	Cyclic peptide end
29	34	82.9	6	23	AAW47559	Cyclic peptide end
30	34	82.9	6	23	AAW47560	Cyclic peptide end
31	34	82.9	7	22	AAW65446	Cyclic peptide wit
32	34	82.9	7	22	AAW65447	Cyclic peptide wit
33	34	82.9	7	23	AAW47556	Cyclic peptide end
34	34	82.9	7	23	AAW47557	Cyclic peptide end
35	34	82.9	7	23	AAW47563	Cyclic peptide end
36	34	82.9	8	23	AAW47562	Cyclic peptide end
37	34	82.9	121	22	AAW28273	Novel human secret
38	34	82.9	198	21	AAW07319	Arabidopsis thalia
39	34	82.9	198	21	AAW07319	Arabidopsis thalia
40	34	82.9	244	21	AAW07318	Arabidopsis thalia
41	34	82.9	244	21	AAW07318	Arabidopsis thalia
42	34	82.9	380	22	ABW67444	Drosophila melanog
43	34	82.9	380	22	ABW67444	Drosophila melanog
44	34	82.9	175	22	AAU50581	Propionibacterium
45	33	80.5	672	22	AAW27202	Parvovirus wild-ty
46	33	80.5	672	22	AAW27204	Parvovirus non-str
47	33	80.5	672	22	AAW27206	Parvovirus non-str
48	33	80.5	672	22	AAW27208	Parvovirus non-str
49	32	78.0	35	22	ABW16761	Human nervous syst
50	32	78.0	57	11	AAW07452	Alzheimer protease
51	32	78.0	64	23	ABP08374	Human ORFX protein
52	32	78.0	80	22	ABW38014	Peptide #5520 enco
53	32	78.0	80	22	ABW38014	Protein #5247 enco
54	32	78.0	80	22	AAW58641	Human brain expres
55	32	78.0	80	22	AAW71147	Human bone marrow
56	32	78.0	80	22	AAW18892	Peptide #5326 enco
57	32	78.0	80	22	AAW31428	Peptide #5465 enco
58	32	78.0	80	23	ABW40944	Human peptide enco
59	32	78.0	90	22	AAU48821	Propionibacterium
60	32	78.0	129	22	AAW62213	Camel ovalbumin bi
61	32	78.0	205	21	AAW88626	H. influenzae dime
62	32	78.0	205	21	AAW88626	H. influenzae dime
63	32	78.0	303	22	AAE00232	Protein encoded by
64	32	78.0	321	22	ABW14801	Novel human diagno
65	32	78.0	354	15	AAW58914	Alpha subunit of h
66	32	78.0	354	22	ABW58951	Drosophila melanog
67	32	78.0	354	22	ABW58951	Drosophila melanog
68	32	78.0	354	22	ABW01175	Novel human diagno
69	32	78.0	354	22	ABW99062	Human G-protein al
70	32	78.0	354	23	ABW09276	G protein-coupled
71	32	78.0	371	22	ABW27689	Novel human diagno
72	32	78.0	390	22	ABW06047	Novel human diagno
73	32	78.0	436	22	ABW27697	Novel human diagno
74	32	78.0	440	22	AAE00249	Taxus cuspidata fu
75	32	78.0	655	22	ABW040911	Novel human diagno

XX AC AAM47521;
 XX DT 12-FEB-2002 (first entry)
 XX DE Cyclic peptide endothelial cell adhesion modulator #14.
 XX KW Cadherin; cytosolic; gynecological; endometriosis;
 XX KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 XX KW cyclic.
 XX OS Synthetic.
 XX PH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Linked to residue 6 to form a cyclic peptide,
 FT N-terminal acetyl or N-terminal CH3-SO2 group"
 FT Modified-site 6
 FT /note= "Linked to residue 1 to form a cyclic peptide,
 FT C-terminal amide"
 XX WO200177146-A2.
 XX 18-OCT-2001.
 XX 09-APR-2001; 2001WO-US11659.
 XX 07-APR-2000; 2000US-0544782.
 XX (UWMC-) UNIV MCGILL.
 XX Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 XX WPI; 2002-049129/06.
 XX Modulating endothelial cell adhesion for inhibiting development of
 XX endometriosis, increasing blood flow to tumor in a mammal, by
 XX contacting cell with a cyclic peptide having cadherin cell adhesion
 XX recognition sequence
 XX Claim 6; Page 64; 139pp; English.
 XX The present invention relates to a method for modulating endothelial cell
 XX adhesion. The method comprises contacting an endothelial cell with a
 XX modulating peptide comprising a cadherin cell adhesion recognition
 XX sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 XX for inhibiting angiogenesis, increasing vasopermeability, increasing
 XX blood flow to a tumour, disrupting neovascularization and inhibiting the
 XX development of endometriosis in a mammal. The modulating peptide reduces
 XX unwanted endothelial adhesion occurring between tumour cells, tumour
 XX cells and normal cells, normal cells as a result of surgery, injury,
 XX chemotherapy, disease and inflammation. The present sequence is one such
 XX modulating peptide.
 XX Sequence 6 AA;
 XX Query Match 100.0%; Score 41; DB 23; Length 6;
 XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 YCHAVC 6
 XX |||||
 XX 1 YCHAVC 6
 XX Db
 XX
 XX RESULT 2
 XX AAY68085
 XX ID AAY68085 standard; Protein; 61 AA.
 XX AC AAY68085;
 XX DT 13-APR-2000 (first entry)
 XX KW

DE Kunitz protease inhibitor variant TW6182.
 XX Kunitz protease inhibitor; KPI; serine protease; kallikrein; plasmin;
 XX amyloid precursor protein; coagulation factor; blood loss; cardiant;
 KW cardiopulmonary bypass surgery; anticoagulant; anti-inflammatory;
 KW anti-arthritic; thrombolytic; antirheumatic; antipsoriasis;
 KW immunosuppressant; pancreatitis; deep vein thrombosis; psoriasis;
 KW rheumatoid arthritis; myocardial infarction; transplant rejection.
 XX Homo sapiens.
 OS Synthetic.
 XX WO9963090-A2.
 XX 09-DEC-1999.
 XX 03-JUN-1999; 99WO-US12276.
 XX 03-JUN-1998; 98US-0087885.
 XX (SCIO-) SCIOS INC.
 XX White RT, Damm D, Lesikar DD, McFadden K, Garrick BL, Lucas AB;
 XX Pollitt NS, Lam AO;
 XX WPI; 2000-105699/09.
 XX Novel enzyme inhibitors especially used to reduce postoperative
 XX bleeding
 XX Example 4; Fig 64; 151pp; English.
 XX The present invention describes protease inhibitors that
 XX are analogues of the Kunitz protease inhibitor (KPI) domain of the
 XX amyloid precursor protein. The protease inhibitors can be used to
 XX treat or prevent disorders associated with increased activity of serine
 XX proteases, specifically blood loss during surgery (particularly
 XX cardiopulmonary bypass surgery where plasma proteases are activated by
 XX contact with surfaces in the heart-lung machine), but also other
 XX conditions such as pancreatitis; deep vein thrombosis; rheumatoid
 XX arthritis; psoriasis; myocardial infarction; and transplant rejection.
 XX They are also for organ preservation and to promote wound healing. In
 XX vitro the protease inhibitors may be used to inhibit serine proteases
 XX during preparation of cell extracts. The protease inhibitors are based
 XX on a human peptide sequence so are unlikely to be immunogenic, and can
 XX produced at high levels in recombinant expression systems, and can
 XX inhibit a wide range of serine proteases. They are more potent or
 XX specific than known inhibitors. The present sequence represents a KPI
 XX variant which is given in an example from the present invention.
 XX Sequence 61 AA;
 XX SQ
 XX Query Match 100.0%; Score 41; DB 21; Length 61;
 XX Best Local Similarity 100.0%; Pred. No. 17;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 YCHAVC 6
 XX |||||
 XX 52 YCHAVC 57
 XX Db
 XX
 XX RESULT 3
 XX AAY68110
 XX ID AAY68110 standard; Protein; 61 AA.
 XX AC AAY68110;
 XX DT 13-APR-2000 (first entry)
 XX DE Kunitz protease inhibitor variant TW6151.
 XX KW Kunitz protease inhibitor; KPI; serine protease; kallikrein; plasmin;
 XX amyloid precursor protein; coagulation factor; blood loss; cardiant;

KW cardiopulmonary bypass surgery; anticoagulant; anti-inflammatory;
 KW anti-arthritis; thrombolytic; antirheumatic; antipsoriatic;
 KW immunosuppressant; pancreatitis; deep vein thrombosis; psoriasis;
 KW rheumatoid arthritis; myocardial infarction; transplant rejection.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX WO963090-A2.
 XX
 PD 09-DEC-1999.
 XX
 PF 03-JUN-1999; 99WO-US12276.
 XX
 PR 03-JUN-1998; 98US-0087885.
 XX
 PA (SCIO-) SCTOS INC.
 XX
 PI White RT, Damm D, Lesikar DD, McFadden K, Garrick BL, Lucas AB;
 PI Pollitt NS, Lam AO;
 XX WPI; 2000-105699/09.
 DR
 XX Novel enzyme inhibitors especially used to reduce postoperative
 PT bleeding -
 PT
 XX Example 4; Fig 64; 151pp; English.
 XX
 CC The present invention describes protease inhibitors that
 CC are analogues of the Kunitz protease inhibitor (KPI) domain of the
 CC amyloid precursor protein. The protease inhibitors can be used used to
 CC treat or prevent disorders associated with increased activity of serine
 CC proteases, specifically blood loss during surgery (particularly
 CC cardiopulmonary bypass surgery where plasma proteases are activated by
 CC contact with surfaces in the heart-lung machine), but also other
 CC conditions such as pancreatitis; deep vein thrombosis; rheumatoid
 CC arthritis; psoriasis; myocardial infarction; and transplant rejection.
 CC They are also for organ preservation and to promote wound healing. In
 CC vitro the protease inhibitors may be used to inhibit serine proteases
 CC during preparation of cell extracts. The protease inhibitors are based
 CC on a human peptide sequence so are unlikely to be immunogenic, can be
 CC produced at high levels in recombinant expression systems, and can
 CC inhibit a wide range of serine proteases. They are more potent or
 CC specific than known inhibitors. The present sequence represents a KPI
 CC variant which is given in an example from the present invention.
 XX
 SQ Sequence 61 AA;
 Query Match 100.0%; Score 41; DB 21; Length 61;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 Db 52 YCHAVC 57
 RESULT 4
 ID AAO09658
 XX AAO09658 standard; Protein; 61 AA.
 AC AAO09658;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 23550.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.

XX WO200164835-A2.
 PN
 XX 07-SEP-2001.
 PD
 XX 26-FEB-2001; 2001WO-US04927.
 XX
 PF 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-514838/56.
 DR N-PSDB; AAI89589.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 PT
 XX Claim 20; SEQ ID NO 23550; 1399pp + Sequence Listing; English.
 PS
 XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 61 AA;
 Query Match 92.7%; Score 38; DB 22; Length 61;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 Db 52 YCHSVC 57
 RESULT 5
 ID AAG98984
 XX AAG98984 standard; Protein; 187 AA.
 AC AAG98984;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE E. coli growth and proliferation related protein sequence SEQ ID NO:454.
 XX
 KW Escherichia coli; growth; proliferation; microbial; antimicrobial;
 KW bacterial infection; microorganism.
 XX
 OS Escherichia coli.
 XX
 PN WO200134810-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US30950.
 PR 09-NOV-1999; 99US-0164415.
 XX
 PA (ELIT-) ELITRA PHARM INC.

XX Forsyth RA, Ohlsen K, Zyskind J;
 XX WPI; 2001-335933/35.
 DR N-PSDB; AAH84655.
 XX Novel nucleic acids that inhibit *Escherichia coli* proliferation, useful
 PT for screening for homologous genes and for designing expression vectors
 PT -
 XX Claim 19; Page 510; 522pp; English.
 PS AAH84373 to AAH84499 represent *Escherichia coli* growth and proliferation
 CC related DNA sequences. (1). AAH84500 to AAH84670 encode the *E. coli*
 CC growth and proliferation related proteins given in AAG99078 and AAG98830
 CC to AAG98999. (1) can be used as potential targets for the generation of
 CC new antimicrobial agents, and for identification of compounds which
 CC interact with the gene products of (1). In addition the expression of
 CC (1) and the purification of the proteins, the purified proteins can be
 CC used to generate reagents and screen small molecule libraries or other
 CC candidate compound libraries for compounds that can be further developed
 CC to yield novel antimicrobial compounds. In addition, nucleic acid probes
 CC complementary to (1) that are specific for particular species of
 CC microorganisms can be used to identify particular microorganism species
 CC in clinical specimens, therefore, providing a rapid and dependable
 CC method by which to identify the causative agents of a bacterial
 CC infection. Also, antibodies generated against proteins translated from
 CC mRNA transcribed from proliferation-required sequences can also be used
 CC to screen for specific microorganisms that produce such proteins in a
 CC species-specific manner. AAH84371 and AAH84670 represent sequencing
 CC primers used in the isolation of *E. coli* growth and proliferation
 CC related sequence, which are used in an example from the present
 CC invention.
 XX SQ Sequence 187 AA;
 Query Match 87.8%; Score 36; DB 22; Length 187;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 Db 67 YCHGIC 72
 RESULT 6
 ID ABB92762
 ID ABB92762 standard; Protein; 428 AA.
 AC ABB92762;
 XX 31-MAY-2002 (first entry)
 XX Herbicidally active polypeptide SEQ ID NO 1973.
 DE Herbicidal; plant; agriculture; herbicide.
 KW Arabidopsis thaliana.
 OS WO200210210-A2.
 PN 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX (FARB) BAYER AG.
 PA Tietjen K, Weidler M;
 PI WPI; 2002-269010/31.
 DR WPI; 2002-269010/31.
 XX

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX Claim 5; SEQ ID NO 1973; 261pp + Sequence Listing; English.
 PS The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX SQ Sequence 428 AA;
 Query Match 87.8%; Score 36; DB 23; Length 428;
 Best Local Similarity 66.7%; Pred. No. 4.9e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 Db 152 FCHGIC 157
 RESULT 7
 ID AAB42468
 ID AAB42468 standard; Protein; 519 AA.
 AC AAB42468;
 XX 08-FEB-2001 (first entry)
 XX Human ORFX ORF2232 polypeptide sequence SEQ ID NO:4464.
 DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; Gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malatia; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX Homo sapiens.
 OS WO200058473-A2.
 PN 05-OCT-2000.
 XX 31-MAR-2000; 2000WO-US08621.
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 XX Shimkets RA, Leach M;
 PI WPI; 2000-602362/57.
 DR N-PSDB; AAC76677.
 XX

PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 3651-3652; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; antitumor; antirheumatic;
 CC antinfiammatory; antibacterial; antiviral; antifungal; antihypertensive;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinfiammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 519 AA;
 SQ

Query Match 87.8%; Score 36; DB 21; Length 519;
 Best Local Similarity 83.3%; Pred. No. 5.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
 |||||
 Db 70 YCHPVC 75

RESULT 8
 AAB61330
 ID AAB61330 standard; protein; 519 AA.
 AC AAB61330;
 XX
 DT 30-MAR-2001 (first entry)
 DE Human transcriptional regulator protein #30.
 XX
 XX Human; transcriptional regulator protein; TXREG.
 KW Homo sapiens.
 OS
 XX WO200078954-A2.
 PN
 XX 28-DEC-2000.
 PD
 XX 15-JUN-2000; 2000WO-US16766.
 PF
 XX 18-JUN-1999; 99US-0140109.
 PR
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B;
 PI WPI; 2001-041425/05.
 DR
 XX Isolated polypeptide with a human transcriptional regulator protein
 PT sequence is useful for the diagnosis, prevention and treatment of
 PT disorders associated with the immune, reproductive and cardiovascular
 PT systems -
 XX
 PS Claim 1; Page 118-119; 142pp; English.

XX The present invention relates to human transcriptional regulator
 CC protein (TXREG) sequences. The antagonist and an agonist of the proteins
 CC of the invention are used to treat disorders associated with decreased
 CC or increased expression or activity of TXREG.

XX Sequence 519 AA;
 SQ

Query Match 87.8%; Score 36; DB 22; Length 519;
 Best Local Similarity 83.3%; Pred. No. 5.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
 |||||
 Db 70 YCHPVC 75

RESULT 9
 AAB42362
 ID AAB42362 standard; Protein; 352 AA.
 XX
 AC AAB42362;
 XX
 DT 08-FEB-2001 (first entry)
 DE Human ORFX ORF2126 polypeptide sequence SEQ ID NO:4252.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinfiammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinfiammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.
 OS
 XX WO200058473-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX 31-MAR-2000; 2000WO-US08621.
 PF
 XX 31-MAR-1999; 99US-0127607.
 PR
 XX 02-APR-1999; 99US-0127636.
 PR
 XX 05-APR-1999; 99US-0127728.
 PR
 XX 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach M;
 PI WPI; 2000-602362/57.
 DR
 XX N-PSDB; AAC76571.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 3451-3452; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; antitumor; antirheumatic;
 CC antinfiammatory; antibacterial; antiviral; antifungal; antihypertensive;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinfiammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predilection to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 352 AA;
 Query Match 85.4%; Score 35; DB 21; Length 352;
 Best Local Similarity 83.3%; Pred. No. 5.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
 DB 196 YCHAPC 201
 |||||

RESULT 10
 AAW23979
 ID AAW23979 standard; peptide; 5 AA.

XX AAW23979;

AC AAW23979;

XX 23-JUL-1998 (first entry)

DE Cadherin-mediated cell adhesion modulating peptide 1.

XX Cadherin-mediated; cell adhesion; drug delivery; treatment; cancer;
 KW angiogenesis; skin grafting; neurite growth; veterinary medicine;
 KW wound healing; tumour; metastasis; carcinoma; leukaemia; modulation;
 KW organ transplant; neurological disease.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT Modified-site 5 /note= "optional N-terminal acetylation"

FT Modified-site 5 /note= "optional C-terminal amidation"

XX WO9802452-A2.

XX 22-JAN-1998.

XX 11-JUL-1997; 97WO-CA00489.

XX 12-JUL-1996; 96US-0021612.

XX (UYMC-) UNIV MCGILL.

XX Blaschuk OW, Gour BJ;

XX WPI; 1998-110522/10.

XX Cyclic peptide(s) that modulate cadherin-mediated cell adhesion -
 PT used to improve drug delivery through skin, to the brain etc., for
 PT treatment of cancer, angiogenesis etc., and to improve grafting of
 PT foreign tissue or neurite growth

PS Claim 8; Page 97; 133pp; English.

XX This is a cyclised peptide capable of modulating cadherin-mediated cell

CC adhesion. Cadherin-expressing cells can be detected from their reaction
 CC with the antibody that binds the peptides of the invention. These
 CC peptides are useful in human or veterinary medicine to modulate adhesion
 CC mediated by E-, N-, P- or R-cadherins (or other cadherins that include
 CC the amino acids HAV recognition sequence), specifically where these are
 CC expressed by epithelial, endothelial, neural or tumour cells or
 CC lymphocytes. The peptides which inhibit cell adhesion are used to improve
 CC delivery of drugs through the skin (such that the peptides enters the
 CC blood stream), to tumours (particularly ovarian or bladder tumours or
 CC melanoma) and to the brain. They are also used to treat cancer
 CC (carcinoma, leukaemia or melanoma), inhibit metastasis and also inhibit
 CC angiogenesis. The peptides that stimulate adhesion are used to improve
 CC wound healing to promote adherence of foreign tissues (skin grafts or
 CC organ transplants), and to improve adherence to tissue culture surfaces
 CC and bioreactors. Other uses of the peptides are to induce apoptosis in
 CC cadherin-expressing cells, increase or decrease neurite outgrowth, to
 CC treat spinal cord injuries and de-myelinating neurological diseases
 CC (specifically multiple sclerosis). They can be used for modulating the
 CC immune system (e.g. in cases of diabetes or rheumatoid arthritis), for
 CC preventing pregnancy and to increase vaso-permeability. The antibodies
 CC which bind to these peptides are also used to modulate cell adhesion and
 CC when coupled to a drug, to target the drug to cadherin-expressing cells.
 CC Diagnostic agents may also be delivered using the peptides in a skin
 CC patch, particularly where the agent generates a colour in contact with,
 CC e.g. cocaine, human immunodeficiency viral proteins, glucose or
 CC prostate-specific antigen, particularly for home-testing kits. The
 CC peptides make it possible to deliver a wide range of drugs through the
 CC skin, avoiding parental administration, by-passing the gastro-
 CC intestinal system and improving patient compliance.

XX Sequence 5 AA;

Query Match 82.9%; Score 34; DB 19; Length 5;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
 DB 1 CHAVC 5
 |||||

RESULT 11

AAV17108

ID AAV17108 standard; peptide; 5 AA.

XX AAV17108;

AC AAV17108;

XX 20-SEP-1999 (first entry)

DE Cadherin-mediated adhesion modulating cyclic peptide. 14

XX Cell adhesion modulation; CAM; synaptic stability; cadherin;
 KW cadherin-mediated adhesion; drug delivery; cell adhesion; tumour;
 KW wound healing; neurite outgrowth; cyclic; circular.

XX Synthetic.

XX WO9933875-A1.

XX 08-JUL-1999.

XX 23-DEC-1998; 98WO-CA01207.

XX 23-DEC-1997; 97US-0996679.

XX (UYMC-) UNIV MCGILL.

XX Blaschuk OW, Gour BJ;

XX WPI; 1999-430231/36.

XX Cyclic peptide cell adhesion modulating agents, useful for
 PT modulating synaptic stability

XX Claim 8; Page 64; 14app; English.

PS The invention provides cyclic peptide cell adhesion modulating (CAM).

CC agents that comprises a His-Ala-Val recognition sequence. Also provided

CC is a method for inhibiting synaptic stability in a mammal that comprises

CC administering to a mammal a therapeutically effective amount of a CAM

CC agent that inhibits cadherin-mediated adhesion, where the agent comprises

CC a cyclic peptide having a peptide ring, and where the sequence His-Ala-

CC Val is present within the peptide ring. The cyclic peptides are cell

CC adhesion modulating agents that inhibit cadherin-mediated adhesion. They

CC can be used in a method for inhibiting synaptic stability in mammals. The

CC agents can be used to treat diseases or other conditions characterized by

CC undesirable cell adhesion or to facilitate drug delivery to a specific

CC tissue or tumour. Alternatively the agents may be used to enhance cell

CC adhesion (e.g. to supplement or replace stitches or to facilitate wound

CC healing) or to enhance or direct neurite outgrowth.

XX

SQ Sequence 5 AA;

Query Match 82.9%; Score 34; DB 20; Length 5;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6

Db 1 CHAVC 5

RESULT 12

AAW91021

ID AAW91021 standard; peptide; 5 AA.

XX

AC AAW91021;

XX

DT 24-MAR-1999 (first entry)

XX

DE Peptide inhibiting interaction of alpha-catenin and beta-catenin.

XX

KW Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;

KW cadherin-mediated function; demyelinating neurological disease;

KW multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;

KW central nervous system; apoptosis induction; cadherin-expression cell;

KW pregnancy prevention; vasopermeability; synaptic stability; diabetes;

KW rheumatoid arthritis; allergic response; learning; memory;

KW antibody-mediated graft rejection.

XX

OS Synthetic.

XX

PN WO9845319-A2.

XX

PD 15-OCT-1998.

XX

PF 14-APR-1998; 98WO-CA00322.

XX

PR 10-APR-1997; 97US-0043361.

XX

PA (UYMC-) UNIV MCGILL.

XX

PI Blaschuk OW, Gour BJ;

XX

DR WPI; 1999-024009/02.

XX

PT New catenin modulating agents - comprising peptides having a

PT sequence HAV or analogues or antibodies; used for modulating

PT cadherin-mediated functions

XX

PS Claim 10; Page 77; 106pp; English.

XX

CC The present sequence represents a peptide which is capable of inhibiting

CC an interaction between alpha-catenin and beta-catenin. The peptide is

CC used in modulating agents that are used for modulating cadherin-mediated

CC functions. They can be used for disrupting interaction between

CC alpha-catenin and beta-catenin in a cell, inhibiting cell adhesion, e.g.

CC between epithelial cells, endothelial cells, neural cells, tumour cells

CC and lymphocytes, for treating a demyelinating neurological disease, e.g.

CC alpha-catenin and beta-catenin in a cell, inhibiting cell adhesion, e.g.

CC between epithelial cells, endothelial cells, neural cells, tumour cells

CC and lymphocytes; for treating a demyelinating neurological disease, e.g.

CC multiple sclerosis, for reducing unwanted cellular adhesion in a mammal,

CC for enhancing the delivery of a drug through the skin of a mammal, for

CC enhancing the delivery of a drug to a tumour in a mammal, for treating

CC cancer in a mammal, for inhibiting angiogenesis in a mammal, for

CC enhancing drug delivery to the central nervous system of a mammal, for

CC inducing apoptosis in a cadherin-expression cell, for modulating the

CC immune system of a mammal, for preventing pregnancy in a mammal, for

CC increasing vasopermeability in a mammal, or for inhibiting synaptic

CC stability in a mammal. In particular they can be used for treating

CC diabetes, rheumatoid arthritis, allergic responses, antibody-mediated

CC graft rejection or for stimulating learning and memory.

XX

SQ Sequence 5 AA;

Query Match 82.9%; Score 34; DB 20; Length 5;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6

Db 1 CHAVC 5

RESULT 13

AAW91033

ID AAW91033 standard; peptide; 5 AA.

XX

AC AAW91033;

XX

DT 24-MAR-1999 (first entry)

XX

DE Peptide inhibiting interaction of alpha-catenin and beta-catenin.

XX

KW Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;

KW cadherin-mediated function; demyelinating neurological disease;

KW multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;

KW central nervous system; apoptosis induction; cadherin-expression cell;

KW pregnancy prevention; vasopermeability; synaptic stability; diabetes;

KW rheumatoid arthritis; allergic response; learning; memory;

KW antibody-mediated graft rejection; cyclic.

XX

OS Synthetic.

XX

PN WO9845319-A2.

XX

PD 15-OCT-1998.

XX

PF 14-APR-1998; 98WO-CA00322.

XX

PR 10-APR-1997; 97US-0043361.

XX

PA (UYMC-) UNIV MCGILL.

XX

PI Blaschuk OW, Gour BJ;

XX

DR WPI; 1999-024009/02.

XX

PT New catenin modulating agents - comprising peptides having a

PT sequence HAV or analogues or antibodies, used for modulating

PT cadherin-mediated functions

XX

PS Claim 12; Page 78; 106pp; English.

XX

CC The present sequence represents a peptide which is capable of inhibiting

CC an interaction between alpha-catenin and beta-catenin. The peptide is

CC used in modulating agents that are used for modulating cadherin-mediated

CC functions. They can be used for disrupting interaction between

CC alpha-catenin and beta-catenin in a cell, inhibiting cell adhesion, e.g.

CC between epithelial cells, endothelial cells, neural cells, tumour cells

CC and lymphocytes, for treating a demyelinating neurological disease, e.g.

CC multiple sclerosis, for reducing unwanted cellular adhesion in a mammal,
 CC for enhancing the delivery of a drug through the skin of a mammal, for
 CC enhancing the delivery of a drug to a tumour in a mammal, for treating
 CC cancer in a mammal, for inhibiting angiogenesis in a mammal, for
 CC enhancing drug delivery to the central nervous system of a mammal, for
 CC inducing apoptosis in a cadherin-expression cell, for modulating the
 CC immune system of a mammal, for preventing pregnancy in a mammal, for
 CC increasing vasopermeability in a mammal, or for inhibiting synaptic
 CC stability in a mammal. In particular they can be used for treating
 CC diabetes, rheumatoid arthritis, allergic responses, antibody-mediated
 CC graft rejection or for stimulating learning and memory.

XX Sequence 5 AA;

Query Match 82.9%; Score 34; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
 DB 1 CHAVC 5

RESULT 14
 AAB27334
 ID AAB27334 standard; Peptide; 5 AA.

XX AAB27334;

XX 15-FEB-2001 (first entry)

DE Beta-catenin derived linear peptide SEQ ID NO: 22.

XX Beta-catenin; cadherin-mediated intercellular adhesion;
 KW cell differentiation; modulating agent; hair loss; skin exfoliation;
 KW internalisation moiety; flanking sequence; transcription; hearing loss.
 XX Unidentified.

XX WO200053632-A1.

XX 14-SEP-2000.

XX 07-MAR-2000; 2000WO-CA00222.

XX 09-MAR-1999; 99US-0265107.

XX (UYMC-) UNIV MCGILL.

XX Blaschuk OW, Gour BJ;

XX WPI; 2000-594308/56.

XX Stimulating beta-catenin mediated gene expression, cellular
 PT differentiation and hair growth, involves contacting cells with
 PT modulating agent capable of inhibiting interaction between alpha and
 PT beta catenin

PS Disclosure; Page 6; 77pp; English.

XX The present invention is concerned with methods of modulating the amount
 CC of free beta-catenin in the cell, and methods of stimulating the
 CC expression of genes involved in cellular differentiation, the
 CC transcription of which is under the control of beta-catenin. The
 CC peptides given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351
 CC can be used as modulating agents which interrupt the interaction between
 CC alpha and beta catenin, causing increased levels of the latter and
 CC stimulating the activation of beta-catenin mediated transcription. This
 CC can be used to stimulate cell differentiation, which can then be used to
 CC promote hair growth and skin exfoliation. This latter is particularly
 CC useful in the improvement of photodamaged skin and to minimise wrinkles.
 CC The modulating peptide can also be used to reduce hearing loss resulting
 CC from inner ear disorders such as hyperacusis and tinnitus.

XX Sequence 5 AA;

Query Match 82.9%; Score 34; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
 DB 1 CHAVC 5

RESULT 15

AAB27346
 ID AAB27346 standard; Peptide; 5 AA.

XX AAB27346;

XX 15-FEB-2001 (first entry)

DE Beta-catenin derived cyclic peptide SEQ ID NO: 34.

XX Beta-catenin; cadherin-mediated intercellular adhesion;
 KW cell differentiation; modulating agent; hair loss; skin exfoliation;
 KW internalisation moiety; flanking sequence; transcription; hearing loss;
 KW cyclic.

XX Unidentified.

XX Key Location/Qualifiers

FT Disulfide-bond 1..5 /note= "the disulfide bond cyclises the molecule"

XX WO200053632-A1.

XX 14-SEP-2000.

XX 07-MAR-2000; 2000WO-CA00222.

XX 09-MAR-1999; 99US-0265107.

XX (UYMC-) UNIV MCGILL.

XX Blaschuk OW, Gour BJ;

XX WPI; 2000-594308/56.

XX Stimulating beta-catenin mediated gene expression, cellular
 PT differentiation and hair growth, involves contacting cells with
 PT modulating agent capable of inhibiting interaction between alpha and
 PT beta catenin

PS Disclosure; Page 7; 77pp; English.

XX The present invention is concerned with methods of modulating the amount
 CC of free beta-catenin in the cell, and methods of stimulating the
 CC expression of genes involved in cellular differentiation, the
 CC transcription of which is under the control of beta-catenin. The
 CC peptides given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351
 CC can be used as modulating agents which interrupt the interaction between
 CC alpha and beta catenin, causing increased levels of the latter and
 CC stimulating the activation of beta-catenin mediated transcription. This
 CC can be used to stimulate cell differentiation, which can then be used to
 CC promote hair growth and skin exfoliation. This latter is particularly
 CC useful in the improvement of photodamaged skin and to minimise wrinkles.
 CC The modulating peptide can also be used to reduce hearing loss resulting
 CC from inner ear disorders such as hyperacusis and tinnitus.

XX Sequence 5 AA;

Query Match 82.9%; Score 34; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 16

AAV73790
ID AAY73790 standard; Peptide; 5 AA.

XX
AC AAY73790;

XX
DT 07-MAR-2000 (first entry)

XX N-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:296.

XX Diagnosis; cancer; OB-cadherin; N-cadherin; E-cadherin; CAR; cyclic;
KW cell adhesion recognition sequence; leukemia; metastasis;
KW prostate cancer; breast cancer; ovarian cancer; carcinoma.

XX Synthetic.

OS Homo sapiens.

XX PN WO9957565-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00362.

XX PR 05-MAY-1998; 98US-0073040.

XX PR 06-NOV-1998; 98US-0187859.

XX PR 20-JAN-1999; 99US-0234395.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Blaschuck OW, Gour BJ, Byers S;

XX DR WPI; 2000-062165/05.

XX Detecting expression of OB-cadherin or N-cadherin used for diagnosing
PT and evaluating cancer

XX PS Disclosure; Page 21; 56pp; English.

XX The present invention describes methods which have been developed for
CC detecting and evaluating cancer that are based on the finding that
CC OB-cadherin and N-cadherin are expressed by metastatic carcinoma cells,
CC but not by highly differentiated, poorly invasive carcinomas. A method
CC for determining the presence or absence of a cancer in a patient
CC comprises: (1) contacting a biological sample from the patient with a
CC binding agent that specifically binds to OB- or N-cadherin OR with an
CC oligonucleotide that hybridizes to a polynucleotide that encodes OB- or
CC N-cadherin; and (2) detecting in the sample an amount of polypeptide
CC that binds to the binding agent OR the amount of polynucleotide that
CC hybridizes to the oligonucleotide, relative to a predetermined cut-off
CC value, and determining the presence or absence of cancer in the patient
CC from this. The methods from the present invention can be used to
CC determine the metastatic potential of a cancer. The methods may be used
CC to detect a metastatic cancer in a patient, to monitor progression of a
CC cancer, or to evaluate the metastatic potential of a cancer. Cancers
CC which may be evaluated using the methods include leukemia, prostate
CC cancer, breast cancer and ovarian cancer. AAY73501 to AAY73813 represent
CC peptide sequences used in the exemplification of the present invention.
CC Specifically, AAY73503 to AAY73789 represent OB-cadherin cell adhesion
CC recognition (CAR) peptides and AAY73790 to AAY73808 represent N-cadherin
CC CAR peptides. AAY47906 to AAY47915 represent PCR primers used in examples
CC from the present invention.

XX Sequence 5 AA;

Query Match 82.9%; Score 34; DB 21; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 17

AAG65370
ID AAG65370 standard; peptide; 5 AA.

XX
AC AAG65370;

XX
DT 30-NOV-2001 (first entry)

XX Cyclic peptide with classical cadherin CAR sequence.

XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
KW cytostatic; vulnerary; immunomodulator; vasotropic; neuroprotective;
KW cerebroprotective; muscular; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

XX FT Modified-site 1

XX FT /note= "N-terminal acetylation"

XX FT Modified-site 5

XX FT /note= "C-terminal amide"

XX PN WO200153331-A2.

XX PD 26-JUL-2001.

XX PF 24-JAN-2001; 2001WO-US02508.

XX PR 24-JAN-2000; 2000US-0491078.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Gour BJ, Blaschuck OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;

XX PI Hu Z;

XX DR WPI; 2001-549899/61.

XX Cell adhesion modulating agent used for enhancing delivery of drug to
PT tumor comprises imidazole compounds

XX PS Claim 15; Page 146; 436pp; English.

XX The invention relates to cell adhesion modulating agents that comprise
CC imidazole compounds of specified formulae that are peptidomimetics of
CC cyclic peptides. The peptidomimetics have a structure similar to that of
CC a cyclic peptide that comprises a cadherin cell recognition sequence RAV.
CC The agents are used for modulating classical cadherin mediated
CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
CC delivery of a drug to a tumor, inhibiting development of cancer,
CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
CC wound healing, modulating the immune system, increasing vasopermeability,
CC treating demyelinating disease, facilitating migration of an N-cadherin
CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
CC neurite outgrowth, and treating spinal cord injuries and macular
CC degeneration. The present sequence represents a cyclic peptide with
CC classical cadherin cell adhesion recognition (CAR) sequence.

XX Sequence 5 AA;

Query Match 82.9%; Score 34; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

2 CHAVC 6
| | | | |
Db 1 CHAVC 5

AC	AAG65441;	
XX	30-NOV-2001 (first entry)	
XX	Cyclic peptide with classical cadherin CAR sequence.	
XX	Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;	
XX	cytostatic; vulnerable; immunomodulator; vasotropic; neuroprotective;	
KW	cerebroprotective; muscular; cyclic.	
XX	Synthetic.	
OS		
XX	Key	Location/Qualifiers
PH	Modified-site 1	/note= "N-terminal acetylation"
FT	Modified-site 6	/note= "C-terminal amide"
FT		
FT		
XX	WO200153331-A2.	
PN		
XX	26-JUL-2001.	
PD		
XX	24-JAN-2001; 2001WO-US02508.	
PF		
XX	24-JAN-2000; 2000US-0491078.	
XX	(ADHE-) ADHEREX TECHNOLOGIES INC.	
PA		
PI	Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;	
PI	Hu Z;	
XX	WPI; 2001-549899/61.	
DR		
XX	Cell adhesion modulating agent used for enhancing delivery of drug to	
PT	tumor comprises imidazole compounds	
PT		
XX	Claim 15; Page 146; 436pp; English.	
PS		
XX	The invention relates to cell adhesion modulating agents that comprise	
CC	imidazole compounds of specified formulae that are peptidomimetics of	
CC	cyclic peptides. The peptidomimetics have a structure similar to that of	
CC	a cyclic peptide that comprises a cadherin cell recognition sequence HAV.	
CC	The agents are used for modulating classical cadherin mediated	
CC	intercellular adhesion, reducing unwanted cellular adhesion, enhancing	
CC	delivery of a drug to a tumor, inhibiting development of cancer,	
CC	inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing	
CC	wound healing, modulating the immune system, increasing vasopermeability,	
CC	treating demyelinating disease, facilitating migration of an N-cadherin	
CC	expressing cell on astrocytes, inhibiting synaptic stability, modulating	
CC	neurite outgrowth, and treating spinal cord injuries and macular	
CC	degeneration. The present sequence represents a cyclic peptide with	
CC	classical cadherin cell adhesion recognition (CAR) sequence.	
XX		
XX	Sequence 6 AA;	
XX		
XX	Query Match 82.9%; Score 34; DB 22; Length 6;	
XX	Best Local Similarity 100.0%; Pred. No. 7.8e+05;	
XX	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	2 CHAVC 6	
DB	1 CHAVC 5	
XX		
XX	RESULT 20	
XX	AAG65444	
XX	ID AAG65444 standard; peptide; 6 AA.	
XX	AAG65444;	
XX		
XX	30-NOV-2001 (first entry)	
DT		
XX	Cyclic peptide with classical cadherin CAR sequence.	

XX KW Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
KW cytotatic; vulnary; immunomodulator; vasotropic; neuroprotective;
XX cerebroprotective; muscular; cyclic.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "putative N-terminal acetylation or
FT alkoxybenzylolation"
FT Modified-site 6 /note= "putative C-terminal amide or ester"
FT WO200153331-A2.
XX 26-JUL-2001.
XX 24-JAN-2001; 2001WO-US02508.
XX 24-JAN-2000; 2000US-0491078.
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
PI Hu Z;
XX WPI; 2001-549899/61.
XX Cell adhesion modulating agent used for enhancing delivery of drug to
PT tumor comprises imidazole compounds -
XX Examples; Page 412; 436pp; English.
XX The invention relates to cell adhesion modulating agents that comprise
CC imidazole compounds of specified formulae that are peptidomimetics of
CC cyclic peptides. The peptidomimetics have a structure similar to that of
CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
CC The agents are used for modulating classical cadherin mediated
CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
CC delivery of a drug to a tumor, inhibiting development of cancer,
CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
CC wound healing, modulating the immune system, increasing vasopermeability,
CC treating demyelinating disease, facilitating migration of an N-cadherin
CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
CC neurite outgrowth, and treating spinal cord injuries and macular
CC degeneration. The present sequence represents a cyclic peptide with
CC classical cadherin cell adhesion recognition (CAR) sequence.
XX Sequence 6 AA;
SQ
Query Match 82.9%; Score 34; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 1 CHAVC 5
RESULT 21
AAG65445
ID AAG65445 standard; peptide; 6 AA.
XX AC AAG65445;
XX 30-NOV-2001 (first entry)
DT
DE Cyclic peptide with classical cadherin CAR sequence.
XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
KW cytotatic; vulnary; immunomodulator; vasotropic; neuroprotective;
KW cerebroprotective; muscular; cyclic.

XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "putative N-terminal acetylation or
FT alkoxybenzylolation"
FT Modified-site 6 /note= "putative C-terminal amide or ester"
FT WO200153331-A2.
XX 26-JUL-2001.
XX 24-JAN-2001; 2001WO-US02508.
XX 24-JAN-2000; 2000US-0491078.
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
PI Hu Z;
XX WPI; 2001-549899/61.
XX Cell adhesion modulating agent used for enhancing delivery of drug to
PT tumor comprises imidazole compounds -
XX Examples; Page 412; 436pp; English.
XX The invention relates to cell adhesion modulating agents that comprise
CC imidazole compounds of specified formulae that are peptidomimetics of
CC cyclic peptides. The peptidomimetics have a structure similar to that of
CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
CC The agents are used for modulating classical cadherin mediated
CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
CC delivery of a drug to a tumor, inhibiting development of cancer,
CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
CC wound healing, modulating the immune system, increasing vasopermeability,
CC treating demyelinating disease, facilitating migration of an N-cadherin
CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
CC neurite outgrowth, and treating spinal cord injuries and macular
CC degeneration. The present sequence represents a cyclic peptide with
CC classical cadherin cell adhesion recognition (CAR) sequence.
XX Sequence 6 AA;
SQ
Query Match 82.9%; Score 34; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 1 CHAVC 5
RESULT 21
AAG65445
ID AAG65445 standard; peptide; 6 AA.
XX AC AAG65445;
XX 30-NOV-2001 (first entry)
DT
DE Cyclic peptide with classical cadherin CAR sequence.
XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
KW cytotatic; vulnary; immunomodulator; vasotropic; neuroprotective;
KW cerebroprotective; muscular; cyclic.

XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "putative N-terminal acetylation or
FT alkoxybenzylolation"
FT Modified-site 6 /note= "putative C-terminal amide or ester"
FT WO200153331-A2.
XX 26-JUL-2001.
XX 24-JAN-2001; 2001WO-US02508.
XX 24-JAN-2000; 2000US-0491078.
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
PI Hu Z;
XX WPI; 2001-549899/61.
XX Cell adhesion modulating agent used for enhancing delivery of drug to
PT tumor comprises imidazole compounds -
XX Examples; Page 412; 436pp; English.
XX The invention relates to cell adhesion modulating agents that comprise
CC imidazole compounds of specified formulae that are peptidomimetics of
CC cyclic peptides. The peptidomimetics have a structure similar to that of
CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
CC The agents are used for modulating classical cadherin mediated
CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
CC delivery of a drug to a tumor, inhibiting development of cancer,
CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
CC wound healing, modulating the immune system, increasing vasopermeability,
CC treating demyelinating disease, facilitating migration of an N-cadherin
CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
CC neurite outgrowth, and treating spinal cord injuries and macular
CC degeneration. The present sequence represents a cyclic peptide with
CC classical cadherin cell adhesion recognition (CAR) sequence.
XX Sequence 6 AA;
SQ
Query Match 82.9%; Score 34; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 2 CHAVC 6
RESULT 22
AAG65448
ID AAG65448 standard; peptide; 6 AA.
XX AC AAG65448;
XX 30-NOV-2001 (first entry)
DT
DE Cyclic peptide with classical cadherin CAR sequence.
XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
KW cytotatic; vulnary; immunomodulator; vasotropic; neuroprotective;
KW cerebroprotective; muscular; cyclic.
XX OS Synthetic.
XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "putative N-terminal acetylation or
FT FT alkoxybenzoylation"
FT Modified-site 6
FT FT /note= "putative C-terminal amide or ester"
XX
XX WO200153331-A2.
XX
XX 26-JUL-2001.
XX
XX 24-JAN-2001; 2001WO-US02508.
XX
XX 24-JAN-2000; 2000US-0491078.
XX
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
XX Hu Z;
XX WPI; 2001-549899/61.
XX
XX Cell adhesion modulating agent used for enhancing delivery of drug to
XX tumor comprises imidazole compounds -
XX
XX Examples; Page 413; 436pp; English.
XX
XX The invention relates to cell adhesion modulating agents that comprise
XX imidazole compounds of specified formulae that are peptidomimetics of
XX cyclic peptides. The peptidomimetics have a structure similar to that of
XX a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
XX The agents are used for modulating classical cadherin mediated
XX intercellular adhesion, reducing unwanted cellular adhesion, enhancing
XX delivery of a drug to a tumor, inhibiting development of cancer,
XX inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
XX wound healing, modulating the immune system, increasing vasopermeability,
XX treating demyelinating disease, facilitating synaptic stability, modulating
XX expressing cell on astrocytes, inhibiting spinal cord injuries and macular
XX degeneration. The present sequence represents a cyclic peptide with
XX classical cadherin cell adhesion recognition (CAR) sequence.
XX Sequence 6 AA;
SQ
Query Match 82.9%; Score 34; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 1 CHAVC 5
RESULT 23
AAG65449
ID AAG65449 standard; peptide; 6 AA.
XX
XX AAG65449;
XX
XX 30-NOV-2001 (first entry)
XX
XX Cyclic peptide with classical cadherin CAR sequence.
XX
XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
XX cytostatic; vulnerable; immunomodulator; vasotropic; neuroprotective;
XX cerebroprotective; muscular; cyclic.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /note= "putative N-terminal acetylation or
XX FT alkoxybenzoylation"
XX
XX Modified-site 6
XX FT /note= "putative C-terminal amide or ester"
XX PN WO200153331-A2.
XX FT

FT
XX
XX WO200153331-A2.
XX
XX 26-JUL-2001.
XX
XX 24-JAN-2001; 2001WO-US02508.
XX
XX 24-JAN-2000; 2000US-0491078.
XX
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
XX Hu Z;
XX WPI; 2001-549899/61.
XX
XX Cell adhesion modulating agent used for enhancing delivery of drug to
XX tumor comprises imidazole compounds -
XX
XX Examples; Page 414; 436pp; English.
XX
XX The invention relates to cell adhesion modulating agents that comprise
XX imidazole compounds of specified formulae that are peptidomimetics of
XX cyclic peptides. The peptidomimetics have a structure similar to that of
XX a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
XX The agents are used for modulating classical cadherin mediated
XX intercellular adhesion, reducing unwanted cellular adhesion, enhancing
XX delivery of a drug to a tumor, inhibiting development of cancer,
XX inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
XX wound healing, modulating the immune system, increasing vasopermeability,
XX treating demyelinating disease, facilitating synaptic stability, modulating
XX expressing cell on astrocytes, inhibiting spinal cord injuries and macular
XX degeneration. The present sequence represents a cyclic peptide with
XX classical cadherin cell adhesion recognition (CAR) sequence.
XX Sequence 6 AA;
SQ
Query Match 82.9%; Score 34; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 1 CHAVC 5
RESULT 24
AAG65450
ID AAG65450 standard; peptide; 6 AA.
XX
XX AAG65450;
XX
XX 30-NOV-2001 (first entry)
XX
XX Cyclic peptide with classical cadherin CAR sequence.
XX
XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
XX cytostatic; vulnerable; immunomodulator; vasotropic; neuroprotective;
XX cerebroprotective; muscular; cyclic.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /note= "putative N-terminal acetylation or
XX FT alkoxybenzoylation"
XX
XX Modified-site 6
XX FT /note= "putative C-terminal amide or ester"
XX PN WO200153331-A2.
XX FT

PD 26-JUL-2001.
 XX
 PF 24-JAN-2001; 2001WO-US02508.
 XX
 PR 24-JAN-2000; 2000US-0491078.
 XX
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 PI Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 PI Hu Z;
 XX
 XX WPI; 2001-549899/61.
 XX
 PT Cell adhesion modulating agent used for enhancing delivery of drug to
 PT tumor comprises imidazole compounds -
 XX
 XX Examples; Page 414; 436pp; English.
 XX
 CC The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cyclic peptide with
 CC classical cadherin cell adhesion recognition (CAR) sequence.
 XX
 SQ Sequence 6 AA;
 Query Match 82.9%; Score 34; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CHAVC 6
 DB 1 CHAVC 5
 RESULT 25
 AAM47551
 ID AAM47551 standard; peptide; 6 AA.
 XX
 AC AAM47551;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Peptide, SEQ ID 84, used to illustrate method for cyclisation.
 XX
 KW Cadherin; cytostatic; gynecological; endometriosis; cyclic;
 KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl. Linked to residue 5 to form a
 FT cyclic peptide"
 FT Modified-site 5 /note= "C-terminal amide. Linked to residue 1 to form a
 FT cyclic peptide"
 XX
 PN WO200177146-A2.
 XX
 PD 18-OCT-2001.
 XX
 XX 09-APR-2001; 2001WO-US11669.
 XX

PR 07-APR-2000; 2000US-0544782.
 XX
 PA (UYMC-) UNIV MCGILL.
 XX
 PI Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 XX
 DR WPI; 2002-049129/06.
 XX
 XX Modulating endothelial cell adhesion for inhibiting development of
 PT endometriosis, increasing blood flow to tumor in a mammal, by
 PT contacting cell with a cyclic peptide having cadherin cell adhesion
 PT recognition sequence -
 XX
 PS Example 1; Page 49; 139pp; English.
 XX
 CC The present invention relates to a method for modulating endothelial cell
 CC adhesion. The method comprises contacting an endothelial cell with a
 CC modulating peptide comprising a cadherin cell adhesion recognition
 CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 CC for inhibiting angiogenesis, increasing vasopermeability, increasing the
 CC blood flow to a tumour, disrupting neovasculation and inhibiting the
 CC development of endometriosis in a mammal. The modulating peptide reduces
 CC unwanted endothelial adhesion occurring between tumour cells, tumour
 CC cells and normal cells, normal cells as a result of surgery, injury, or
 CC chemotherapy, disease and inflammation. The present sequence was used to
 CC illustrate the present invention.
 XX
 SQ Sequence 6 AA;
 Query Match 82.9%; Score 34; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CHAVC 6
 DB 1 CHAVC 5
 Search completed: June 20, 2003, 20:51:50
 Job time : 40.3182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 12.6818 Seconds
(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-10
Perfect score: 41
Sequence: 1 YCHAVC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	6	US-09-458-870-100	Sequence 100, Appl
2	34	82.9	5	US-08-893-534A-8	Sequence 8, Appli
3	34	82.9	5	US-08-996-679-8	Sequence 8, Appli
4	34	82.9	5	US-09-115-395-10	Sequence 10, Appl
5	34	82.9	5	US-09-507-102-8	Sequence 8, Appli
6	34	82.9	5	US-09-250-059-10	Sequence 10, Appl
7	34	82.9	5	US-09-248-074-10	Sequence 10, Appl
8	34	82.9	5	US-09-357-717-10	Sequence 10, Appl
9	34	82.9	5	US-09-458-870-10	Sequence 10, Appl
10	34	82.9	6	US-09-458-870-84	Sequence 84, Appl
11	34	82.9	6	US-09-458-870-87	Sequence 87, Appl
12	34	82.9	6	US-09-458-870-88	Sequence 88, Appl
13	34	82.9	6	US-09-458-870-91	Sequence 91, Appl
14	34	82.9	6	US-09-458-870-92	Sequence 92, Appl
15	34	82.9	6	US-09-458-870-93	Sequence 93, Appl
16	34	82.9	7	US-09-458-870-89	Sequence 89, Appl
17	34	82.9	7	US-09-458-870-90	Sequence 90, Appl
18	34	82.9	7	US-09-458-870-96	Sequence 96, Appl
19	34	82.9	8	US-09-458-870-95	Sequence 95, Appl
20	34	82.9	8	US-09-458-870-101	Sequence 101, Appl
21	32	78.0	205	US-08-549-515-6	Sequence 6, Appli
22	32	78.0	205	US-08-549-515-12	Sequence 12, Appl
23	32	78.0	207	US-08-549-515-9	Sequence 9, Appli
24	32	78.0	303	US-09-457-046B-12	Sequence 12, Appl
25	32	78.0	440	US-09-457-046B-45	Sequence 45, Appl
26	31	75.6	15	5496552-13	Patent No. 5496552
27	31	75.6	19	US-08-566-800A-18	Sequence 18, Appl

28	31	75.6	19	1	US-08-566-800A-19	Sequence 19, Appl
29	31	75.6	19	1	US-08-206-310A-23	Sequence 23, Appl
30	31	75.6	19	1	US-08-206-310A-24	Sequence 24, Appl
31	31	75.6	19	2	US-08-398-010A-23	Sequence 23, Appl
32	31	75.6	19	2	US-08-398-010A-24	Sequence 24, Appl
33	31	75.6	19	2	US-08-398-628A-23	Sequence 23, Appl
34	31	75.6	19	2	US-08-398-628A-24	Sequence 24, Appl
35	31	75.6	19	2	US-08-932-589-18	Sequence 18, Appl
36	31	75.6	19	2	US-08-932-589-19	Sequence 19, Appl
37	31	75.6	19	2	US-08-399-115A-23	Sequence 23, Appl
38	31	75.6	19	2	US-08-399-115A-24	Sequence 24, Appl
39	31	75.6	24	1	US-08-463-155A-23	Sequence 23, Appl
40	31	75.6	24	1	US-08-463-155A-24	Sequence 24, Appl
41	31	75.6	24	1	US-08-463-432B-23	Sequence 23, Appl
42	31	75.6	24	1	US-08-463-432B-24	Sequence 24, Appl
43	31	75.6	52	4	US-09-369-494-11	Sequence 11, Appl
44	31	75.6	55	4	US-09-369-494-3	Sequence 3, Appli
45	31	75.6	55	4	US-09-358-569D-3	Sequence 3, Appli
46	31	75.6	56	1	US-08-155-331-17	Sequence 17, Appl
47	31	75.6	56	1	US-08-424-022-17	Sequence 17, Appl
48	31	75.6	56	2	US-08-424-017B-17	Sequence 17, Appl
49	31	75.6	56	5	PCT-US93-11696-17	Sequence 17, Appl
50	31	75.6	56	6	5223482-30	Patent No. 5223482
51	31	75.6	57	2	US-08-422-333-11	Sequence 11, Appl
52	31	75.6	57	2	US-08-829-876-1	Sequence 1, Appli
53	31	75.6	57	2	US-08-829-876-3	Sequence 3, Appli
54	31	75.6	57	2	US-08-829-876-114	Sequence 114, App
55	31	75.6	57	2	US-08-829-876-116	Sequence 116, App
56	31	75.6	57	2	US-08-829-876-128	Sequence 128, App
57	31	75.6	57	2	US-08-829-876-152	Sequence 152, App
58	31	75.6	57	2	US-08-829-876-162	Sequence 162, App
59	31	75.6	57	2	US-08-829-876-165	Sequence 165, App
60	31	75.6	57	4	US-09-234-874A-1	Sequence 1, Appli
61	31	75.6	57	4	US-09-234-874A-3	Sequence 3, Appli
62	31	75.6	57	4	US-09-234-874A-5	Sequence 5, Appli
63	31	75.6	57	4	US-09-234-874A-114	Sequence 114, App
64	31	75.6	57	4	US-09-234-874A-116	Sequence 116, App
65	31	75.6	57	4	US-09-234-874A-128	Sequence 128, App
66	31	75.6	57	4	US-09-234-874A-152	Sequence 152, App
67	31	75.6	57	4	US-09-234-874A-162	Sequence 162, App
68	31	75.6	57	4	US-09-234-874A-165	Sequence 165, App
69	31	75.6	57	6	5187153-19	Patent No. 5187153
70	31	75.6	57	6	5220013-21	Patent No. 5220013
71	31	75.6	57	6	5220013-27	Patent No. 5220013
72	31	75.6	57	6	5223482-21	Patent No. 5223482
73	31	75.6	58	1	US-08-384-489-12	Sequence 12, Appl
74	31	75.6	58	1	US-08-358-160-73	Sequence 73, Appl
75	31	75.6	58	1	US-08-566-800A-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-458-870-100
; Sequence 100, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-458-870-100

Query Match 100.0%; Score 41; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YHAVC 6
| | | | |
Db 1 YHAVC 6

RESULT 2
US-08-893-534A-8
Sequence 8, Application US/08893534A
Patent No. 6031072
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
CELL ADHESION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue.
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,534A
FILING DATE: 11-JUL-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-893-534A-8

Query Match 82.9%; Score 34; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 3
US-08-996-679-8
Sequence 8, Application US/08996679
Patent No. 6169071

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
CELL ADHESION
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-8

Query Match 82.9%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 4
US-09-115-395-10
Sequence 10, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115,395A
CURRENT FILING DATE: 1998-07-14
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/996,679
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID-NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis

US-09-115-395-10

Query Match 82.9%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 5

US-09-507-102-8
; Sequence 8, Application US/09507102
; Patent No. 6326352

GENERAL INFORMATION:

APPLICANT: Blaschuk, Orest W.

Gour, Barbara J.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESS: SEED IP LAW GROUP PLLC
STREET: 6300 Bank of America Bldg., 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/507,102

FILING DATE: 17-Feb-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/893,534

FILING DATE: 11-JUL-1997

APPLICATION NUMBER: US 60/021,612

FILING DATE: 12-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Christiansen, William T.

REGISTRATION NUMBER: 44,614

REFERENCE/DOCKET NUMBER: 100086.401C10

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

SEQUENCE CHARACTERISTICS:

INFORMATION FOR SEQ ID NO: 8:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: circular

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-507-102-8

Query Match 82.9%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 6

US-09-250-059-10

; Sequence 10, Application US/09250059

; Patent No. 633307

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH

; FILE REFERENCE: 100086.401C5

; CURRENT APPLICATION NUMBER: US/09/250,059

; CURRENT FILING DATE: 1999-02-12

; NUMBER OF SEQ ID NOS: 87

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide

; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence

; FEATURE:

; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal

; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group

; OTHER INFORMATION: and/or C-terminal modifications such as amide or

; OTHER INFORMATION: ester group

US-09-250-059-10

Query Match 82.9%; Score 34; DB 4; Length 5;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6

Db 1 CHAVC 5

RESULT 7

US-09-248-074-10

; Sequence 10, Application US/09248074

; Patent No. 6346512

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION

; FILE REFERENCE: 100086.401C5

; CURRENT APPLICATION NUMBER: US/09/248,074

; CURRENT FILING DATE: 1999-02-10

; NUMBER OF SEQ ID NOS: 81

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide

; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence

; FEATURE:

; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal

; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group

; OTHER INFORMATION: and/or C-terminal modifications such as amide or

; OTHER INFORMATION: ester group

US-09-248-074-10

Query Match 82.9%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 8

US-09-357-717-10

; Sequence 10, Application US/09357717

; Patent No. 6417325

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-357-717-10

Query Match 82.9%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 9
US-09-458-870-10
; Sequence 10, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-10

Query Match 82.9%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 10
US-09-458-870-84
; Sequence 84, Application US/09458870
; Patent No. 6465427

; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-84

Query Match 82.9%; Score 34; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 11
US-09-458-870-87
; Sequence 87, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-87

Query Match 82.9%; Score 34; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 12

US-09-458-870-88
; Sequence 88, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-88
Query Match 82.9%; Score 34; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 2 CHAVC 6
RESULT 13
US-09-458-870-91
; Sequence 91, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-91
Query Match 82.9%; Score 34; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 1 CHAVC 5
RESULT 14
US-09-458-870-92
; Sequence 92, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-92
Query Match 82.9%; Score 34; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 1 CHAVC 5
RESULT 15
US-09-458-870-93
; Sequence 93, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 93
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-93
Query Match 82.9%; Score 34; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 1 CHAVC 5

QY 2 CHAVC 6
|||||
Db 1 CHAVC 5

RESULT 16
US-09-458-870-89
; Sequence 89, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-89

Query Match 82.9%; Score 34; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
|||||
Db 1 CHAVC 5

RESULT 17
US-09-458-870-90
; Sequence 90, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-90

Query Match 82.9%; Score 34; DB 4; Length 7;

Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
|||||
Db 2 CHAVC 6

RESULT 18
US-09-458-870-96
; Sequence 96, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 96
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-96

Query Match 82.9%; Score 34; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
|||||
Db 3 CHAVC 7

RESULT 19
US-09-458-870-95
; Sequence 95, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-95

Query Match 82.9%; Score 34; DB 4; Length 7;

US-09-458-870-95

Query Match 82.9%; Score 34; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
|||
DB 3 CHAVC 7

RESULT 20

US-09-458-870-101
; Sequence 101, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: Where Xaa is beta,beta-dimethyl cysteine
US-09-458-870-101

Query Match 82.9%; Score 34; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
|||
DB 3 CHAVC 7

RESULT 21

US-08-549-515-6
; Sequence 6, Application US/08549515
; Patent No. 6054123
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: Haemophilus Influenzae
; TITLE OF INVENTION: Dimethylsulphoxide Reductase Enzyme
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 Unviersity Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
US-09-458-870-95

OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/549,515
; FILING DATE: 27-OCT-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-522
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-549-515-6

Query Match 78.0%; Score 32; DB 3; Length 205;
Best Local Similarity 66.7%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||
DB 104 YCHMAC 109

RESULT 22

US-08-549-515-12
; Sequence 12, Application US/08549515
; Patent No. 6054123
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: Haemophilus Influenzae
; TITLE OF INVENTION: Dimethylsulphoxide Reductase Enzyme
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 Unviersity Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/549,515
; FILING DATE: 27-OCT-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-522
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-549-515-12

Query Match 78.0%; Score 32; DB 3; Length 205;
Best Local Similarity 66.7%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCHAVC 6
||| |
Db 104 YCHMAC 109

RESULT 23

US-08-549-515-9
; Sequence 9, Application US/08549515
; Patent No. 6054123
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: Haemophilus Influenzae
; TITLE OF INVENTION: Dimethylsulphoxide Reductase Enzyme
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/549,515
; FILING DATE: 27-OCT-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-522
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-549-515-9

Query Match 78.0%; Score 32; DB 3; Length 207;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCHAVC 6
||| |
Db 104 YCHMAC 109

RESULT 24

US-09-457-046B-12
; Sequence 12, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12.
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Taxus cuspidata

US-09-457-046B-12

Query Match 78.0%; Score 32; DB 4; Length 303;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
:|:|:
Db 86 FCHGIC 91

RESULT 25

US-09-457-046B-45
; Sequence 45, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-457-046B-45

Query Match 78.0%; Score 32; DB 4; Length 440;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
:|:|:
Db 160 FCHGIC 165

Search completed: June 20, 2003, 21:00:27
Job time : 13.6818 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:52:19 ; Search time 15.4091 seconds
(without alignments)
42.134 Million cell updates/sec

Title: US-10-105-008-10

Perfect score: 41

Sequence: 1 YCHAVC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pap:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pap:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pap:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pap:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pap:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pap:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pap:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	82.9	5	9	US-09-769-145-10
2	34	82.9	5	9	US-10-058-821-10
3	34	82.9	5	10	US-09-234-395-296
4	34	82.9	5	10	US-09-305-928-296
5	34	82.9	5	12	US-10-006-982-8
6	34	82.9	6	9	US-09-769-145-81
7	34	82.9	6	9	US-09-769-145-84
8	34	82.9	6	9	US-09-769-145-85
9	34	82.9	6	9	US-09-769-145-88
10	34	82.9	6	9	US-09-769-145-89
11	34	82.9	6	9	US-09-769-145-90
12	34	82.9	7	9	US-09-769-145-86
13	34	82.9	7	9	US-09-769-145-87
14	34	82.9	380	12	US-10-003-806-8
15	34	82.9	537	9	US-10-047-542-74
16	32	78.0	73	9	US-10-083-357-1221
17	32	78.0	80	10	US-09-864-761-38546
18	32	78.0	303	9	US-09-866-570A-12
19	32	78.0	303	9	US-10-166-984-12

20	78.0	303	10	US-09-866-572A-12	Sequence 12, Appl
21	78.0	354	9	US-09-952-680A-22	Sequence 22, Appl
22	78.0	440	9	US-09-866-570A-45	Sequence 45, Appl
23	78.0	440	9	US-10-166-984-45	Sequence 45, Appl
24	78.0	440	10	US-09-866-572A-45	Sequence 45, Appl
25	78.0	758	10	US-09-996-194-13	Sequence 13, Appl
26	75.6	20	10	US-09-864-761-35523	Sequence 35523, A
27	75.6	55	10	US-09-864-761-33588	Sequence 33588, A
28	75.6	56	10	US-09-864-761-37763	Sequence 37763, A
29	75.6	63	10	US-09-864-761-34873	Sequence 34873, A
30	75.6	67	9	US-10-154-333-29	Sequence 29, Appl
31	75.6	96	9	US-09-798-518-2	Sequence 2, Appl
32	75.6	96	9	US-09-798-518-3	Sequence 3, Appl
33	75.6	101	1	US-08-260-675-9	Sequence 9, Appl
34	75.6	102	1	US-08-260-675-11	Sequence 11, Appl
35	75.6	117	9	US-10-154-333-7	Sequence 7, Appl
36	75.6	117	9	US-10-154-333-15	Sequence 15, Appl
37	75.6	117	12	US-10-115-406-5	Sequence 5, Appl
38	75.6	117	12	US-10-115-406-13	Sequence 13, Appl
39	75.6	118	9	US-09-859-211-37	Sequence 37, Appl
40	75.6	118	9	US-09-880-708-12	Sequence 12, Appl
41	75.6	118	9	US-09-880-708-15	Sequence 15, Appl
42	75.6	118	9	US-09-872-856-37	Sequence 37, Appl
43	75.6	118	9	US-10-154-333-17	Sequence 8, Appl
44	75.6	118	10	US-09-813-459-8	Sequence 10, Appl
45	75.6	118	10	US-09-813-459-10	Sequence 10, Appl
46	75.6	118	12	US-10-115-406-15	Sequence 15, Appl
47	75.6	129	9	US-09-804-625-2	Sequence 2, Appl
48	75.6	143	12	US-10-002-278-3	Sequence 3, Appl
49	75.6	143	12	US-10-002-278-4	Sequence 4, Appl
50	75.6	172	1	US-08-957-425-17	Sequence 17, Appl
51	75.6	184	9	US-09-791-932-106	Sequence 106, Appl
52	75.6	214	10	US-09-995-515-5	Sequence 5, Appl
53	75.6	215	9	US-09-791-279-145	Sequence 145, Appl
54	75.6	364	9	US-10-066-500-65	Sequence 65, Appl
55	75.6	364	9	US-10-028-072-342	Sequence 342, Appl
56	75.6	364	9	US-10-121-049-342	Sequence 342, Appl
57	75.6	364	9	US-10-123-904-342	Sequence 342, Appl
58	75.6	364	9	US-10-140-470-342	Sequence 342, Appl
59	75.6	364	9	US-10-175-746-342	Sequence 342, Appl
60	75.6	364	9	US-10-176-918-342	Sequence 342, Appl
61	75.6	364	9	US-10-176-921-342	Sequence 342, Appl
62	75.6	364	9	US-10-002-796-65	Sequence 65, Appl
63	75.6	364	9	US-10-066-273-65	Sequence 65, Appl
64	75.6	364	9	US-10-066-494-65	Sequence 65, Appl
65	75.6	364	9	US-10-137-865-342	Sequence 342, Appl
66	75.6	364	9	US-10-140-474-342	Sequence 342, Appl
67	75.6	364	9	US-10-142-431-342	Sequence 342, Appl
68	75.6	364	9	US-10-143-114-342	Sequence 342, Appl
69	75.6	364	9	US-10-066-269-65	Sequence 65, Appl
70	75.6	364	9	US-10-140-002-342	Sequence 342, Appl
71	75.6	364	9	US-10-066-193-65	Sequence 65, Appl
72	75.6	364	9	US-10-066-211-65	Sequence 65, Appl
73	75.6	364	9	US-10-142-419-342	Sequence 342, Appl
74	75.6	364	9	US-10-123-262-342	Sequence 342, Appl
75	75.6	364	9	US-10-142-423-342	Sequence 342, Appl

ALIGNMENTS

RESULT 1
US-09-769-145-10
; Sequence 10, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Ammar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoameng

APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-10

Query Match 82.9%; Score 34; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 2
US-10-058-821-10
Sequence 10, Application US/10058821
Publication No. US20030087811A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
FILE REFERENCE: 100086.401C12
CURRENT APPLICATION NUMBER: US/10/058,821
CURRENT FILING DATE: 2002-01-29
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-10-058-821-10

Query Match 82.9%; Score 34; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 3
US-09-234-395-296
Sequence 296, Application US/09234395
Patent No. US20020123044A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.

APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 296
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-234-395-296

Query Match 82.9%; Score 34; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 4
US-09-305-928-296
Sequence 296, Application US/09305928
Patent No. US20020146687A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C4
CURRENT APPLICATION NUMBER: US/09/305,928
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 296
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-305-928-296

Query Match 82.9%; Score 34; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 5
US-10-006-982-8
Sequence 8, Application US/10006982
Patent No. US20020151475A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
CELL ADHESION

NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED IP LAW GROUP PLLC
STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/006,982
FILING DATE: 04-Dec-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Christensen, William T.
REGISTRATION NUMBER: 44,614
REFERENCE/DOCKET NUMBER: 100086.401C11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-006-982-8

Query Match 82.9%; Score 34; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
DB 1 CHAVC 5

RESULT 6
US-09-769-145-81
; Sequence 81, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Ammar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoameng
; APPLICANT: Hu, Zengjian
; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
; FILE REFERENCE: 100086.413C1
; CURRENT APPLICATION NUMBER: US/09/769,145
; CURRENT FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 09/491,078
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: With Classical Cell Adhesion Recognition Sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: modification such as amide or

; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-769-145-81

Query Match 82.9%; Score 34; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
DB 1 CHAVC 5

RESULT 7
US-09-769-145-84
; Sequence 84, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Ammar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoameng
; APPLICANT: Hu, Zengjian
; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
; FILE REFERENCE: 100086.413C1
; CURRENT APPLICATION NUMBER: US/09/769,145
; CURRENT FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 09/491,078
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: With Classical Cell Adhesion Recognition Sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-769-145-84

Query Match 82.9%; Score 34; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
DB 1 CHAVC 5

RESULT 8
US-09-769-145-85
; Sequence 85, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Ammar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoameng
; APPLICANT: Hu, Zengjian
; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
; FILE REFERENCE: 100086.413C1
; CURRENT APPLICATION NUMBER: US/09/769,145
; CURRENT FILING DATE: 2001-01-24

;; PRIOR APPLICATION NUMBER: US 09/491,078
;; PRIOR FILING DATE: 2000-01-24
;; NUMBER OF SEQ ID NOS: 96
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 85
;; LENGTH: 6
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
;; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
;; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
;; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
;; OTHER INFORMATION: and/or C-terminal modifications such as amide or
;; OTHER INFORMATION: ester group
US-09-769-145-85

Query Match 82.9%; Score 34; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 2 CHAVC 6

RESULT 9

US-09-769-145-88
;; Sequence 88, Application US/09769145
;; Patent No. US20020168761A1
;; GENERAL INFORMATION:
;; APPLICANT: Gour, Barbara J.
;; APPLICANT: Blaschuk, Orest W.
;; APPLICANT: Ali, Annmar
;; APPLICANT: Ni, Feng
;; APPLICANT: Chen, Zhigang
;; APPLICANT: Michaud, Stephanie
;; APPLICANT: Wang, Shoameng
;; APPLICANT: Hu, Zengjian
;; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
;; FILE REFERENCE: 100086.413C1
;; CURRENT APPLICATION NUMBER: US/09/769,145
;; CURRENT FILING DATE: 2001-01-24
;; PRIOR APPLICATION NUMBER: US 09/491,078
;; PRIOR FILING DATE: 2000-01-24
;; NUMBER OF SEQ ID NOS: 96
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 88
;; LENGTH: 6
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
;; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
;; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
;; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
;; OTHER INFORMATION: and/or C-terminal modifications such as amide or
;; OTHER INFORMATION: ester group
US-09-769-145-88

Query Match 82.9%; Score 34; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 10

US-09-769-145-89
;; Sequence 89, Application US/09769145
;; Patent No. US20020168761A1

;; GENERAL INFORMATION:
;; APPLICANT: Gour, Barbara J.
;; APPLICANT: Blaschuk, Orest W.
;; APPLICANT: Ali, Annmar
;; APPLICANT: Ni, Feng
;; APPLICANT: Chen, Zhigang
;; APPLICANT: Michaud, Stephanie
;; APPLICANT: Wang, Shoameng
;; APPLICANT: Hu, Zengjian
;; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
;; FILE REFERENCE: 100086.413C1
;; CURRENT APPLICATION NUMBER: US/09/769,145
;; CURRENT FILING DATE: 2001-01-24
;; PRIOR APPLICATION NUMBER: US 09/491,078
;; PRIOR FILING DATE: 2000-01-24
;; NUMBER OF SEQ ID NOS: 96
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 89
;; LENGTH: 6
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
;; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
;; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
;; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
;; OTHER INFORMATION: and/or C-terminal modifications such as amide or
;; OTHER INFORMATION: ester group
US-09-769-145-89

Query Match 82.9%; Score 34; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHAVC 6
Db 1 CHAVC 5

RESULT 11

US-09-769-145-90
;; Sequence 90, Application US/09769145
;; Patent No. US20020168761A1
;; GENERAL INFORMATION:
;; APPLICANT: Gour, Barbara J.
;; APPLICANT: Blaschuk, Orest W.
;; APPLICANT: Ali, Annmar
;; APPLICANT: Ni, Feng
;; APPLICANT: Chen, Zhigang
;; APPLICANT: Michaud, Stephanie
;; APPLICANT: Wang, Shoameng
;; APPLICANT: Hu, Zengjian
;; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
;; FILE REFERENCE: 100086.413C1
;; CURRENT APPLICATION NUMBER: US/09/769,145
;; CURRENT FILING DATE: 2001-01-24
;; PRIOR APPLICATION NUMBER: US 09/491,078
;; PRIOR FILING DATE: 2000-01-24
;; NUMBER OF SEQ ID NOS: 96
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 90
;; LENGTH: 6
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
;; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
;; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
;; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
;; OTHER INFORMATION: and/or C-terminal modifications such as amide or
;; OTHER INFORMATION: ester group
US-09-769-145-90

Query Match 82.9%; Score 34; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
| | | | |
Db 1. CHAVC 5

RESULT 12

US-09-769-145-86

; Sequence 86, Application US/09769145

; Patent No. US20020168761A1

; GENERAL INFORMATION:

; APPLICANT: Gour, Barbara J.

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Ali, Anmar

; APPLICANT: Ni, Feng

; APPLICANT: Chen, Zhigang

; APPLICANT: Michaud, Stephanie

; APPLICANT: Wang, Shoameng

; APPLICANT: Hu, Zengjian

; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION

; FILE REFERENCE: 100086.413C1

; CURRENT APPLICATION NUMBER: US/09/769,145

; CURRENT FILING DATE: 2001-01-24

; PRIOR APPLICATION NUMBER: US 09/491,078

; PRIOR FILING DATE: 2000-01-24

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 86

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide

; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence

; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal

; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group

; OTHER INFORMATION: and/or C-terminal modifications such as amide or

; OTHER INFORMATION: ester group

US-09-769-145-86

Query Match 82.9%; Score 34; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
| | | | |
Db 1 CHAVC 5

RESULT 13

US-09-769-145-87

; Sequence 87, Application US/09769145

; Patent No. US20020168761A1

; GENERAL INFORMATION:

; APPLICANT: Gour, Barbara J.

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Ali, Anmar

; APPLICANT: Ni, Feng

; APPLICANT: Chen, Zhigang

; APPLICANT: Michaud, Stephanie

; APPLICANT: Wang, Shoameng

; APPLICANT: Hu, Zengjian

; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION

; FILE REFERENCE: 100086.413C1

; CURRENT APPLICATION NUMBER: US/09/769,145

; CURRENT FILING DATE: 2001-01-24

; PRIOR APPLICATION NUMBER: US 09/491,078

; PRIOR FILING DATE: 2000-01-24

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 87

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide

; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence

; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal

; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group

; OTHER INFORMATION: and/or C-terminal modifications such as amide or

; OTHER INFORMATION: ester group

US-09-769-145-87

Query Match 82.9%; Score 34; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
| | | | |
Db 2 CHAVC 6

RESULT 14

US-10-003-806-8

; Sequence 8, Application US/10003806

; Patent No. US20020119929A1

; GENERAL INFORMATION:

; APPLICANT: Bishop, Colin E.

; APPLICANT: Agoulnik, Alexander I.

; APPLICANT: Zhu, Qichao

; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY

; FILE REFERENCE: P02066US1/10024824

; CURRENT APPLICATION NUMBER: US/10/003,806

; CURRENT FILING DATE: 2001-11-02

; PRIOR APPLICATION NUMBER: 60/245,872

; PRIOR FILING DATE: 2000-11-03

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 380

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-10-003-806-8

Query Match 82.9%; Score 34; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
| | | | |
Db 318 CHAVC 322

RESULT 15

US-10-047-542-74

; Sequence 74, Application US/10047542

; Patent No. US20020168367A1

; GENERAL INFORMATION:

; APPLICANT: LARRICK, JAMES W.

; APPLICANT: WYCOFF, KEITH L.

; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL

; TITLE OF INVENTION: AND BACTERIAL DISEASES

; FILE REFERENCE: 030905.0004.CIP1

; CURRENT APPLICATION NUMBER: US/10/047,542

; CURRENT FILING DATE: 2001-10-26

; PRIOR APPLICATION NUMBER: PCT/US01/13932

; PRIOR FILING DATE: 2001-04-28

; PRIOR APPLICATION NUMBER: 60/200,298

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 101

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 74

; LENGTH: 537

; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-047-542-74

Query Match 82.9%; Score 34; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
Db 96 CHAVC 100

RESULT 16
US-10-083-357-1221
; Sequence 1221, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiangdong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 1221
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1221

Query Match 78.0%; Score 32; DB 9; Length 73;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
Db 46 FCHKVC 51

RESULT 17
US-09-864-761-38546
; Sequence 38546, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aconica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38546
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006070.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
US-09-864-761-38546

Query Match 78.0%; Score 32; DB 10; Length 80;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCHAVC 6
Db 42 YCHPTC 47

RESULT 18
US-09-866-570A-12
; Sequence 12, Application US/09866570A
; Patent No. US20020168745A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/866,570A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US/09/457,046B
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-866-570A-12

Query Match 78.0%; Score 32; DB 9; Length 303;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
Db 86 FCHGIC 91

RESULT 19

```
US-10-166-984-12
; Sequence 12, Application US/10166984
; Publication No. US20030108891A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 4630-62267
; CURRENT APPLICATION NUMBER: US/10/166,984
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/866,570
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/457,046
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 09/411,145
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-166-984-12

Query Match      78.0%; Score 32; DB 9; Length 303;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 YCHAVC 6
      :||:|
Db      86 FCHGIC 91

RESULT 20
US-09-866-572A-12
; Sequence 12, Application US/09866572A
; Patent No. US20020138859A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/866,572A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/457,046
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-866-572A-12

Query Match      78.0%; Score 32; DB 9; Length 303;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 YCHAVC 6
      :||:|
Db      86 FCHGIC 91

US-10-166-984-45
; Sequence 45, Application US/10166984
; Publication No. US20030108891A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 4630-62267
; CURRENT APPLICATION NUMBER: US/10/166,984
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/866,570
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/457,046
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 09/411,145
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-166-984-45

Query Match      78.0%; Score 32; DB 9; Length 440;
Best Local Similarity 50.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 YCHAVC 6
      :||:|
Db      160 FCHGIC 165

RESULT 23
US-10-166-984-45
; Sequence 45, Application US/10166984
; Publication No. US20030108891A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 4630-62267
; CURRENT APPLICATION NUMBER: US/10/166,984
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/866,570
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/457,046
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 09/411,145
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-166-984-45

Query Match      78.0%; Score 32; DB 9; Length 440;
Best Local Similarity 50.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 YCHAVC 6
      :||:|
Db      160 FCHGIC 165

US-10-166-984-45
; Sequence 45, Application US/09866570A
; Patent No. US20020168745A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/866,570A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US/09/457,046B
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-866-570A-45

Query Match      78.0%; Score 32; DB 9; Length 354;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 YCHAVC 6
      :||:|
Db      320 YCHWTC 325

RESULT 22
US-09-866-570A-45
; Sequence 45, Application US/09866570A
; Patent No. US20020168745A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/866,570A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US/09/457,046B
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-866-570A-45

Query Match      78.0%; Score 32; DB 9; Length 354;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 YCHAVC 6
      :||:|
Db      320 YCHWTC 325

RESULT 23
US-10-166-984-45
; Sequence 45, Application US/10166984
; Publication No. US20030108891A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 4630-62267
; CURRENT APPLICATION NUMBER: US/10/166,984
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/866,570
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/457,046
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 09/411,145
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-166-984-45

Query Match      78.0%; Score 32; DB 9; Length 440;
```

Best Local Similarity 50.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6,
:|:|:
Db 160 FCHGIC 165

RESULT 24

US-09-866-572A-45
; Sequence 45, Application US/09866572A
; Patent No. US20020138859A1
; GENERAL INFORMATION:
; APPLICANT: Coreau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/866,572A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/457,046
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 45
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-866-572A-45

Query Match 78.0%; Score 32; DB 10; Length 440;
Best Local Similarity 50.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
:|:|:
Db 160 FCHGIC 165

RESULT 25

US-09-996-194-13
; Sequence 13, Application US/09996194
; Patent No. US20020151696A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhara
; TITLE OF INVENTION: 84242, 8035, 55304, 52999, and 21999, and Methods of Use Thereof
; FILE REFERENCE: 35800/240590
; CURRENT APPLICATION NUMBER: US/09/996,194
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/250,348
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,073
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/253,878
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,338
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 758
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-996-194-13

Query Match 78.0%; Score 32; DB 10; Length 758;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAV 5
:|:|:
Db 123 YCHAV 127

Search completed: June 20, 2003, 21:02:29
Job time : 16.4091 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 15.2727 Seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-10
Perfect score: 41
Sequence: 1 YHAVC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match.0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	36	87.8	187 1 Q0EC30	prepilin peptidase
2	36	87.8	187 2 B90809	prepilin peptidase
3	36	87.8	187 2 D85934	prepilin peptidase
4	36	87.8	428 2 T48008	hypothetical prote
5	34	82.9	76 2 H64393	hypothetical prote
6	34	82.9	244 2 E84885	hypothetical prote
7	34	82.9	1111 2 T26972	hypothetical prote
8	33	80.5	179 2 T25841	hypothetical prote
9	33	80.5	374 2 S69627	hypothetical prote
10	33	80.5	402 1 MNXRPC	nonstructural prote
11	33	80.5	402 2 S35639	hypothetical prote
12	33	80.5	421 2 A12090	two-component resp
13	33	80.5	433 2 A39556	triacylglycerol li
14	33	80.5	575 2 T48579	auxin-regulated pr
15	33	80.5	578 2 H86365	T26J12.7 protein -
16	33	80.5	668 1 A44276	noncapsid protein
17	33	80.5	672 1 UYPV11	noncapsid protein
18	33	80.5	672 1 UYPV1M	noncapsid protein
19	33	80.5	721 1 UYPV1M	noncapsid protein
20	32	78.0	82 2 T17912	hypothetical prote
21	32	78.0	157 2 T17883	major capsid prote
22	32	78.0	205 2 F64109	dimethylsulfoxide
23	32	78.0	205 2 A06612	anaerobic dimethyl
24	32	78.0	205 2 AG0680	probable dimethyl
25	32	78.0	205 2 B85615	hypothetical prote
26	32	78.0	205 2 D85764	probable oxidoredu
27	32	78.0	205 2 D90751	hypothetical prote
28	32	78.0	205 2 G90915	probable oxidoredu
29	32	78.0	205 2 AH0403	anaerobic dimethyl

30	32	78.0	205	2 F64828	anaerobic dimethyl
31	32	78.0	205	2 G64914	dimethylsulfoxide
32	32	78.0	354	1 RGHU01	GRP-binding regula
33	32	78.0	354	1 RGHU01	GRP-binding regula
34	32	78.0	354	1 RGHU01	GRP-binding regula
35	32	78.0	354	1 RGHU01	GRP-binding regula
36	32	78.0	354	1 RGHU01	GRP-binding regula
37	32	78.0	354	1 RGHU01	GRP-binding regula
38	32	78.0	354	1 RGHU01	GRP-binding regula
39	32	78.0	354	1 RGHU01	GRP-binding regula
40	32	78.0	354	2 A61035	hypothetical prote
41	32	78.0	354	2 T19476	hypothetical prote
42	32	78.0	440	2 T52320	10-deacetylacacati
43	32	78.0	461	2 D87971	protein Y438A.2 [
44	32	78.0	466	2 H71314	probable K+ transp
45	32	78.0	502	2 T26851	hypothetical prote
46	32	78.0	530	2 B96616	hypothetical prote
47	32	78.0	590	2 T20537	hypothetical prote
48	32	78.0	610	2 T16761	hypothetical prote
49	32	78.0	760	2 F86281	protein F1086.14 [
50	32	78.0	1481	1 Q2D0P3	pyrimidine synthe
51	31	75.6	47	2 B36826	oncogenic protein
52	31	75.6	76	2 S04855	Alzheimer's diseas
53	31	75.6	76	2 S03607	Alzheimer's diseas
54	31	75.6	76	2 S06678	Alzheimer's diseas
55	31	75.6	96	2 S35715	glycine-rich prote
56	31	75.6	100	2 A32282	Alzheimer's diseas
57	31	75.6	111	2 S41082	amyloid precursor
58	31	75.6	157	2 T18512	hypothetical prote
59	31	75.6	161	2 T22277	hypothetical prote
60	31	75.6	162	2 T33127	hypothetical prote
61	31	75.6	171	2 S09903	hypothetical prote
62	31	75.6	171	2 S09759	hypothetical prote
63	31	75.6	203	2 T33126	hypothetical prote
64	31	75.6	231	2 QJ7702	NT-3 antigen - hum
65	31	75.6	243	2 S34794	osmotin - common t
66	31	75.6	250	2 C97812	hypothetical prote
67	31	75.6	251	2 E86378	protein F2109.10 [
68	31	75.6	288	2 T00862	probable serine/th
69	31	75.6	307	1 YXMS7	thymidylate syntha
70	31	75.6	307	2 S53715	thymidylate syntha
71	31	75.6	313	1 YXMS7	thymidylate syntha
72	31	75.6	328	2 A05299	fibrinogen beta ch
73	31	75.6	351	2 A34201	bone morphogenetic
74	31	75.6	352	2 T24154	hypothetical prote
75	31	75.6	352	2 T47820	hypothetical prote

ALIGNMENTS

RESULT 1

Q0EC30
prepilin peptidase dependent protein B precursor - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 31-Mar-1998 #sequence_revision 05-Dec-1997 #text_change 01-Mar-2002
C:Accession: B55065; C24137
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B55065
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-187 <BLAT>
A:Cross-references: GB:U00096; NID:g1789185; PIDN:AAC75864.1; PID:g178911
A:Experimental source: strain K-12, substrain M61655
R:Finch, P.W.; Wilson, R.E.; Brown, K.; Hickson, I.D.; Tomkinson, A.E.; Emerson, P.T
Nucleic Acids Res. 14, 4437-4451, 1986
A:Title: Complete nucleotide sequence of the Escherichia coli recC gene and of the th
A:Reference number: A93625; MUID:96232583; PMID:3520484
A:Accession: C24137

A;Molecule type: DNA
A;Residues: MVPCFRWGYLSWQFTIGLCATLARSNERDTEACFLNPAQYRNGRAYSLQKLNRRVVGFGPAVKTPALSA
A;Cross-references: GB:X03966; NID:G42684; PID:CAA27601.1; PID:G42686
C;Genetics:
A;Gene: ppdB
A;Map position: 61 min
C;Superfamily: prepilin peptidase dependent protein B precursor

Query Match 87.8%; Score 36; DB 1; Length 187;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
67 YCHGIC 72

Db

RESULT 2
B98089
prepilin peptidase dependent protein B [imported] - Escherichia coli (strain O157:H7, su
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: B98089
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A98629; MUID:21156231; PMID:11258796
A;Accession: B98089
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-187 <HAY>
A;Cross-references: GB:BA000007; PID:BA37105.1; PID:g13363154; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs3682
C;Superfamily: prepilin peptidase dependent protein B precursor

Query Match 87.8%; Score 36; DB 2; Length 187;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
67 YCHGIC 72

Db

RESULT 3
D85934
Prepilin peptidase dependent protein B [imported] - Escherichia coli (strain O157:H7, su
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: D85934
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85934
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-187 <STO>
A;Cross-references: GB:AE005174; NID:g12517310; PIDN:AAG57936.1; GSPDB:GN00145; UMGF:Z41
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ppdB
C;Superfamily: prepilin peptidase dependent protein B precursor

Query Match 87.8%; Score 36; DB 2; Length 187;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
67 YCHGIC 72

Db

us-10-105-008-10.rpr

Db 67 YCHGIC 72
|||:|
67 YCHGIC 72

RESULT 4
T48008
hypothetical protein T17J13.120 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T48008
R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Maye
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24482
A;Accession: T48008
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-428 <RIE>
A;Cross-references: EMBL:AL138651
A;Experimental source: cultivar Columbia; BAC clone T17J13
C;Genetics:
A;Map position: 3
A;Introns: 137/3
A;Note: T17J13.120

Query Match 87.8%; Score 36; DB 2; Length 428;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
152 FCHATC 157

Db

RESULT 5
H64393
hypothetical protein MJ0752 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Oct-1999
C;Accession: H64393
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A;Reference number: A64300; MUID:96337999; PMID:8686087
A;Accession: H64393
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-76 <BUL>
A;Cross-references: GB:U67521; GB:L77117; NID:g1591463; PIDN:AAB98754.1; PID:g1499571
C;Genetics:
A;Map position: REV677944-677714

Query Match 82.9%; Score 34; DB 2; Length 76;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
65 YCHSAC 70

Db

RESULT 6
E84885
hypothetical protein At2g45010 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: E84885
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
eug, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: E84885
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <STO>
 A:Cross-references: GB:AE002093; NID:g4895249; PIDN:AAD32834.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g45010
 A:Map position: 2

Query Match 82.9%; Score 34; DB 2; Length 244;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
 |||||
 Db 106 CHAVC 110

RESULT 7

T26972 hypothetical protein Y47H9C.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T26972

R:Harris, B.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293

A:Accession: T26972

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1111 <WIL>

A:Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GN00019; CESP:Y47H9C.4

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:Y47H9C.4

A:Map position: 1

A:Introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match 82.9%; Score 34; DB 2; Length 1111;

Best Local Similarity 66.7%; Pred. No. 3e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6

|||||

Db 847 FCHALC 852

RESULT 8

T25841 hypothetical protein M03F4.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T25841

R:Fulton, L.

submitted to the EMBL Data Library, July 1996

A:Description: The sequence of C. elegans cosmid M03F4.

A:Reference number: Z20097

A:Accession: T25841

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-179 <FUL>

A:Cross-references: EMBL:U64601; PIDN:AAB04577.1; GSPDB:GN00028; CESP:M03F4.6

A:Experimental source: strain Bristol N2; clone M03F4

C:Genetics:

A:Gene: CESP:M03F4.6

A:Map position: X

A:Introns: 41/1; 129/2

Query Match 80.5%; Score 33; DB 2; Length 179;

Best Local Similarity 66.7%; Pred. No. 1.3e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 |||||
 Db 37 FCHATC 42

RESULT 9

S69627

hypothetical protein YDR459C - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C>Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 19-Apr-2002

C:Accession: S69627

R:Dietrich, F.S.

submitted to the EMBL Data Library, August 1995

A:Description: The sequence of S. cerevisiae cosmids 9410, 8035, 8166, and 9787.

A:Reference number: S69554

A:Accession: S69627

A:Molecule type: DNA

A:Residues: 1-374 <DIE>

A:Cross-references: EMBL:U33050; NID:g927726; PID:g927749; GSPDB:GN00004; MIPS:YDR459

C:Genetics:

A:Gene: MIPS:YDR459C

A:Cross-references: SGD:S0002867

A:Map position: 4R

Query Match 80.5%; Score 33; DB 2; Length 374;

Best Local Similarity 66.7%; Pred. No. 2.1e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6

|||||

Db 34 YCHKLC 39

RESULT 10

MXRPC

nonstructural protein NS34 - porcine rotavirus C (strain Cowden)

C:Species: porcine rotavirus C

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: A41040

R:Qian, Y.; Jiang, B.; Saif, L.J.; Kang, S.Y.; Ojeh, C.K.; Green, K.Y.

VIology 184, 752-757, 1991

A:Title: Molecular analysis of the gene 6 from a porcine group C rotavirus that encod

A:Reference number: A41040; MUID:91361567; PMID:1653496

A:Accession: A41040

A:Molecule type: genomic RNA

A:Residues: 1-402 <QIA>

A:Cross-references: GB:M69115; NID:g333307; PIDN:AAA47087.1; PID:g333308

C:Genetics:

A:Map position: segment 6

C:Superfamily: rotavirus nonstructural protein

C:Keywords: nonstructural protein

Query Match 80.5%; Score 33; DB 1; Length 402;

Best Local Similarity 80.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

|||||

Db 280 CHAIC 284

RESULT 11

S35639

hypothetical protein - bovine rotavirus C

C:Species: bovine rotavirus C

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999

C:Accession: S35639

R:Jiang, B.; Tsunemitsu, H.; Gentsch, J.R.; Saif, L.J.; Glass, R.I.

Nucleic Acids Res. 21, 2250, 1993

A:Title: Nucleotide sequences of genes 6 and 10 of a bovine group C rotavirus.

A:Reference number: S35639; MUID:93275758; PMID:8389040

A;Accession: S35639
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-402 <JIA>
A;Cross-references: EMBL:L12390
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1993
C;Superfamily: rotavirus nonstructural protein

Query Match 80.5%; Score 33; DB 2; Length 402;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YCHAVC 6
|||:|
DB 280 CHAIC 284

RESULT 12

A12090
two-component response/regulator alr2280 (Imported) - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: A12090
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: A12090
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-421 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA073979.1; PID:gl17131371; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr2280

Query Match 80.5%; Score 33; DB 2; Length 421;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
DB 340 YCNAIC 345

RESULT 13

A39556
triacylglycerol lipase (EC 3.1.1.3) 2 - Moraxella sp. (strain TA144)
C;Species: Moraxella sp.
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 08-Oct-1999
C;Accession: A39556
R;Feller, G.; Thiry, M.; Garday, C.
DNA Cell Biol. 10, 381-388, 1991
A;Title: Nucleotide sequence of the lipase gene lip2 from the antarctic psychrotroph Moraxella
A;Reference number: A39556; MUID:91321741; PMID:1907455
A;Accession: A39556
A;Molecule type: DNA
A;Residues: 1-433 <FEL>
A;Cross-references: GB:X53868; NID:G296841; PIDN:CAA37862.1; PID:G296842
A;Gene: lip2
C;Genetics:
C;Keywords: carboxylic ester hydrolase

Query Match 80.5%; Score 33; DB 2; Length 433;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
DB 180 FCHTVC 185

RESULT 14

T48579
auxin-regulated protein GH3 homolog T31B5.140 - Arabidopsis thaliana
N;Alternate names: protein T31B5.140
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 18-Jul-2001
C;Accession: T48579
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24490
A;Accession: T48579
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-575 <BEV>
A;Cross-references: EMBL:AL163491
A;Experimental source: cultivar Columbia; BAC clone T31B5
C;Genetics:
A;Map position: 5
A;Intons: 95/2; 129/2; 390/1
A;Note: T31B5.140
C;Superfamily: soybean auxin-regulated protein GH3

Query Match 80.5%; Score 33; DB 2; Length 575;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
DB 198 YCHLLC 203

RESULT 15

H86365
T2GJ12.7 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001
C;Accession: H86365
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alorain, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, D.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Maris, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talbot, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-578 <STO>
A;Cross-references: GB:AE005172; NID:G2829896; PIDN:AAC00604.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: soybean auxin-regulated protein GH3

Query Match 80.5%; Score 33; DB 2; Length 578;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
|||:|
DB 201 YCHLLC 206

RESULT 16

A44276
noncapsid protein NS1 - parvovirus Lu111
C;Species: parvovirus Lu111
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999
C;Accession: A44276

R;Diffoot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993

A;Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A;Reference number: A44276; MUID:93297126; PMID:8517025

A;Accession: A44276

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-668 <DIF>

A;Cross-references: GB:M81888

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 80.5%; Score 33; DB 1; Length 668;

Best Local Similarity 80.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

Db 378 CHAIC 382

RESULT 17

UYPVIM

noncapsid protein NS1 - parvovirus H1

C;Species: parvovirus H1

A;Note: host Homo sapiens (man)

C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 16-Jul-1999

C;Accession: A03695

R;Rhode III, S.L.; Paradiso, P.R.

J. Virol. 45, 173-184, 1983

A;Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybridization

A;Reference number: A03695; MUID:83112183; PMID:6823009

A;Accession: A03695

A;Molecule type: DNA

A;Residues: 1-672 <RHO>

A;Cross-references: EMBL:X01457; NID:G60993; PIDN:CAA25689.1; PID:G60994; EMBL:J02198

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 80.5%; Score 33; DB 1; Length 672;

Best Local Similarity 80.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

Db 378 CHAIC 382

RESULT 18

UYPVIM

noncapsid protein NS1 - minute virus of mice

C;Species: minute virus of mice, murine parvovirus

C;Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994

C;Accession: A03696

R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.

Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.

A;Reference number: A03696; MUID:83143341; PMID:6298737

A;Accession: A03696

A;Molecule type: DNA

A;Residues: 1-672 <AST>

A;Cross-references: EMBL:V01115

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 80.5%; Score 33; DB 1; Length 672;

Best Local Similarity 80.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

Db 378 CHAIC 382

RESULT 19

UYPVIM

noncapsid protein NS1 - minute virus of mice (strain MM1)

C;Species: minute virus of mice, murine parvovirus

C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C;Accession: A23008; A29510

R;Sahli, R.; McMaster, G.K.; Hirt, B.

Nucleic Acids Res. 13, 3617-3633, 1985

A;Title: DNA sequence comparison between two tissue-specific variants of the autonomous

A;Reference number: A23008; MUID:85242059; PMID:3855242

A;Accession: A23008

A;Molecule type: DNA

A;Residues: 1-721 <SAH>

A;Cross-references: EMBL:X02481

R;Astell, C.R.; Gardiner, E.M.; Tattersall, P.

J. Virol. 57, 656-669, 1986

A;Title: DNA sequence of the lymphotropic variant of minute virus of mice, MM(1), a

A;Reference number: A29510; MUID:86115415; PMID:3502703

A;Accession: A29510

A;Molecule type: DNA

A;Residues: 1-645, 1', 647-721 <AST>

A;Cross-references: EMBL:M12032; NID:G332289; PIDN:AAA69566.1; PID:G825477

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 80.5%; Score 33; DB 1; Length 721;

Best Local Similarity 80.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

Db 427 CHAIC 431

RESULT 20

T17912

hypothetical protein a409R - Chlorella virus PBCV-1

C;Species: Chlorella virus PBCV-1

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T17912

R;Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A;Reference number: Z18806

A;Accession: T17912

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-82 <GRA>

A;Cross-references: EMBL:U42580; NID:G4028896; PIDN:AAC96777.1

A;Experimental source: specific host Chlorella strain NC64

C;Genetics:

A;Note: a409R

Query Match

Best Local Similarity 78.0%; Score 32; DB 2; Length 82;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCHAVC 6

Db 44 YCHMAC 49

RESULT 21

T17883

major capsid protein homolog A382R - Chlorella virus PBCV-1

C;Species: Chlorella virus PBCV-1

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T17883

R;Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A;Reference number: Z18806

A;Accession: T17883

A;Status: preliminary; translated from GB/EMBL/DBJ

Mon Jun 23 10:08:15 2003

us-10-105-008-10.rpr

Page 7

Search completed: June 20, 2003, 20:58:54
Job time : 18.2727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 ; Search time 7.22727 Seconds
(without alignments)
34.433 Million cell updates/sec

Title: US-10-105-008-10

Perfect score: 41

Sequence: 1 YCHAVC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	87.8	187	1 PPDB_ECOLI	P08371 escherichia
2	34	82.9	76	1 Y52 METJA	Q58162 methanococc
3	33	80.5	402	1 V534 ROTBS	P34717 bovine rota
4	33	80.5	402	1 V534 ROTPC	P27586 porcine rot
5	33	80.5	433	1 LIP2_MORSP	P24484 moraxella s
6	33	80.5	451	1 TEG_ENTHI	P54401 entamoeba h
7	33	80.5	668	1 VNC5_PAVL3	P36311 parvovirus
8	33	80.5	672	1 VNC5_MUMIV	P07300 murine minu
9	33	80.5	672	1 VNC5_PAVHH	P03134 hamster par
10	33	80.5	672	1 VNC5_PAVHH	P03133 hamster par
11	32	78.0	204	1 DMSB_ECOLI	P18776 escherichia
12	32	78.0	204	1 YNFG_ECOLI	P77313 escherichia
13	32	78.0	205	1 DMSB_HAEIN	P45003 haemophilus
14	32	78.0	353	1 GB01_BOVIN	P08239 bos taurus
15	32	78.0	353	1 GB01_HUMAN	P09471 homo sapien
16	32	78.0	353	1 GB01_MOUSE	P18872 mus musculu
17	32	78.0	353	1 GB01_RAT	P04900 rattus norv
18	32	78.0	353	1 GB0_CABEL	P51875 caenorhabdi
19	32	78.0	353	1 GB0_HELTI	P51877 helisoma tr
20	32	78.0	353	1 GB0_LYMSI	P30683 lymsiae sta
21	32	78.0	354	1 GB01_DROME	P16378 drosophila
22	32	78.0	354	1 GB01_LOCOM	P38404 locusta mig
23	32	78.0	356	1 GB0_PATYE	P15976 patinopecte
24	32	78.0	592	1 TORZ_CABEL	P98060 caenorhabdi
25	32	78.0	778	1 SZ1B_BRARE	Q9w686 brachydanio
26	32	78.0	2185	1 PYR1_DICDI	P20054 dictyosteli
27	31	75.6	76	1 A4_MACNU	P29216 macaca mula
28	31	75.6	87	1 A4_MACFA	P53601 macaca fasc
29	31	75.6	96	1 GRP7_DAUCA	P37704 daucus caro
30	31	75.6	171	1 IR10_HCMVA	P16808 human cytom
31	31	75.6	307	1 TYSY_MOUSE	P07607 mus musculu
32	31	75.6	307	1 TYSY_RAT	P45352 rattus norv
33	31	75.6	312	1 TYSY_HUMAN	P04818 homo sapien

34	31	75.6	352	1 GBA7_CABEL	Q21917 caenorhabdi
35	31	75.6	353	1 BMP2_CHICK	Q90751 gallus gall
36	31	75.6	355	1 DVR1_BRARE	P35621 brachydanio
37	31	75.6	357	1 GBA1_CABEL	P28051 caenorhabdi
38	31	75.6	364	1 GDF3_HUMAN	Q9nr23 homo sapien
39	31	75.6	366	1 GDF3_MOUSE	Q07104 mus musculu
40	31	75.6	366	1 IHA_MOUSE	Q04997 mus musculu
41	31	75.6	366	1 IHA_RAT	P17490 rattus norv
42	31	75.6	393	1 BMP2_RAT	P49001 rattus norv
43	31	75.6	394	1 BMP2_MOUSE	P21274 mus musculu
44	31	75.6	395	1 BMP2_RABIT	O45564 oryctolagus
45	31	75.6	396	1 BMP2_DAMDA	O19006 dama dama (
46	31	75.6	396	1 BMP2_HUMAN	P12643 homo sapien
47	31	75.6	398	1 BMPA_XENLA	P25703 xenopus lae
48	31	75.6	398	1 BMPB_XENLA	P30884 xenopus lae
49	31	75.6	479	1 FIBB_RAT	P14480 rattus norv
50	31	75.6	588	1 DECA_DROME	P07713 drosophila
51	31	75.6	593	1 DECA_DROSI	P91706 drosophila
52	31	75.6	621	1 DECA_DROPS	P91699 drosophila
53	31	75.6	707	1 GCVK_HCMVA	P16788 human cytom
54	31	75.6	751	1 A4_SHISC	Q95241 salmizi sci
55	31	75.6	763	1 APP2_HUMAN	Q05481 homo sapien
56	31	75.6	765	1 APP2_RAT	P15943 rattus norv
57	31	75.6	770	1 A4_HUMAN	P05067 homo sapien
58	31	75.6	770	1 A4_MOUSE	P12023 mus musculu
59	31	75.6	770	1 A4_RAT	P08592 rattus norv
60	31	75.6	1210	1 EGFR_HUMAN	P00533 homo sapien
61	31	75.6	1305	1 DPOG_NEUCR	Q9v767 neurospora
62	30	73.2	144	1 GRP1_CHERU	P11898 chenopodium
63	30	73.2	144	1 GRP9_DAUCA	P37703 daucus caro
64	30	73.2	184	1 INAI_HORSE	P50003 equus cabal
65	30	73.2	208	1 CRP2_HUMAN	P52943 homo sapien
66	30	73.2	208	1 CRP2_RAT	P36201 rattus norv
67	30	73.2	225	1 NRFC_HAEIN	P45015 haemophilus
68	30	73.2	265	1 IHA_SHEEP	P38440 ovis aries
69	30	73.2	285	1 VIE1_HSVB4	P27426 bovine herp
70	30	73.2	329	1 IHA_CHICK	P43031 gallus gall
71	30	73.2	360	1 IHA_BOVIN	P07755 trichosurus
72	30	73.2	361	1 IHA_TRIVU	O77755 trichosurus
73	30	73.2	364	1 IHA_PIG	P04087 eus scrofa
74	30	73.2	366	1 IHA_HUMAN	P05111 homo sapien
75	30	73.2	367	1 IHA_HORSE	P55101 equus cabal

ALIGNMENTS

RESULT 1
PPDB_ECOLI STANDARD; PRT; 187 AA.
ID PPDB_ECOLI
AC P08371;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prepilin peptidase dependent protein B precursor.
GN PPDB OR B2825.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86232583; PubMed=3520484;
RA Finch P.W., Wilson R.E., Brown K., Hickson I.D., Tomkinson A.E.,
RA Emmerson P.T.;
RT "Complete nucleotide sequence of the Escherichia coli recC gene and
of the thvA-recC intergenic region.";
RL Nucleic Acids Res. 14:4437-4451(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.O.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1453-1474 (1997).
RN [3]
RP IDENTIFICATION
RX MEDLINE=95020523; PubMed=7934814;
RA Hobbs M., Mattick J.S.;
RT "Common components in the assembly of type 4 fimbriae, DNA transfer
RT systems, filamentous phage and protein-secretion apparatus: a general
RT system for the formation of surface-associated protein complexes.";
RL Mol. Microbiol. 10:233-243 (1993).
CC -1- FUNCTION: NOT YET KNOWN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X03966; CAA27601.1; ALT INIT.
CC EMBL; U29581; AAB40472.1; -;
CC EMBL; AE000366; AAC75864.1; -;
CC PIR; C24137; QQEC30.
CC EcoGene; EG1156; pppB.
CC InterPro; IPR001120; Prok N methyltrn.
CC PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
CC Methylation; Complete proteome.
FT PROPEP 1 BY SIMILARITY.
FT CHAIN 8 PREPILIN PEPTIDASE DEPENDENT PROTEIN B.
FT MOD RES 8 8 METHYLATION (BY SIMILARITY).
FT MOD RES 8 8 METHYLATION (BY SIMILARITY).
SQ SEQUENCE 187 AA; 20519 MW; 433303D846D786F0 CRC64;
Query Match 87.8%; Score 36; DB 1; Length 187;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 YCHAVC 6
DB 67 YCHGIC 72
RESULT 2
ID Y752 METJA STANDARD; PRT; 76 AA.
AC Q58162;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0752.
GN MJ0752.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8680807;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
RT jannaschii.";
RL Science 273:1058-1073 (1996).

CC -1- SIMILARITY: TO THE N-TERMINAL OF M. JANNASCHII MJ0857.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67521; AAB98754.1; -;
CC TIGR; MJ0752; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 76 AA; 9186 MW; ABDC436DE6EC4F8 CRC64;
Query Match 82.9%; Score 34; DB 1; Length 76;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 YCHAVC 6
DB 65 YCHSAC 70
RESULT 3
ID VN34 ROTBS STANDARD; PRT; 402 AA.
AC P34717;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nonstructural RNA-binding protein 34 (NS34) (NCVP4).
GN S6.
OS Bovine rotavirus (group C / strain Shintoku).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=33723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93275758; PubMed=8389040;
RA Jiang B., Tsunemitsu H., Gentsch J.R., Saif L.J., Glass R.I.;
RT "Nucleotide sequences of genes 6 and 10 of a bovine group C
RT rotavirus.";
RL Nucleic Acids Res. 21:2250-2250 (1993).
CC -1- SIMILARITY: CONTAINS 1 DREM (DOUBLE-STRANDED RNA-BINDING) DOMAIN
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L12390; -; NOT_ANNOTATED_CS.
CC PIR; S35639; S35639.
CC InterPro; IPR001159; DS_RBD.
CC InterPro; IPR002873; Rota_NSP3.
CC Pfam; PF00035; dsrm; 1.
CC Pfam; PF01665; Rota_NSP3; 1.
CC SMART; SM00358; DSRM; 1.
CC PROSITE; PS00137; DS_RBD; 1.
CC Nonstructural protein; RNA-binding.
FT DOMAIN 384 400 DREM.
SQ SEQUENCE 402 AA; 45081 MW; EBF553B9D95D49 CRC64;
Query Match 80.5%; Score 33; DB 1; Length 402;
Best Local Similarity 80.0%; Pred. No. 67;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 280 CHAIC 284


```

RESULT 4
VN34 ROTPC
ID_VN34_ROTPC STANDARD; PRT; 402 AA.
AC P27586;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nonstructural RNA-binding protein 34 (NS34) (NCVP4).
GN S6.
OS Porcine rotavirus (group C / strain Cowden).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=10916;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91361587; PubMed=1653496;
RA Qian Y.A., Jiang B.M., Saif L.J., Kang S.Y., Ojeh C.K., Green K.Y.;
RT "Molecular analysis of the gene 6 from a porcine group C rotavirus
that encodes the NS34 equivalent of group A rotaviruses.";
RL Virology 184:752-757(1991).
CC -1- SIMILARITY: CONTAINS 1 DBRM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M69115; AAA47087.1; -.
DR PIR; A41040; MNXRPC.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR002873; RotA_NSP3.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF01665; RotA_NSP3; 1.
DR PROSITE; PS50137; DS_RBD; 1.
KW Nonstructural protein; RNA-binding.
FT DOMAIN 384 400 DBRM.
SQ SEQUENCE 402 AA; 45125 MW; 8B11F7AC7FB5135C CRC64;

Query Match 80.5%; Score 33; DB 1; Length 402;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
DB 280 CHAIC 284

RESULT 5
LIP2_MORSP
ID_LIP2_MORSP STANDARD; PRT; 433 AA.
AC P24484;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Lipase 2 (EC 3.1.1.13) (Triacylglycerol lipase).
GN Lip2 OR L2.
OS Moraxella sp.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae; Moraxella.
OX NCBI_TaxID=479;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RC STRAIN=TA144;
RX MEDLINE=91321741; PubMed=1907455;
RA Feller G., Thiry M., Gerday C.;
RT "Nucleotide sequence of the lipase gene lip2 from the antarctic
psychrotroph Moraxella TA144 and site-specific mutagenesis of the
conserved serine and histidine residues.";
RL DNA Cell Biol. 10:381-388(1991).
CC -1- FUNCTION: CATALYZES LIPOLYSIS AT TEMPERATURES AS LOW AS 3 DEGREES
CELSIUS.
CC

```

```

CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
fatty acid anion.
CC -1- SIMILARITY: BELONGS TO THE "GDXG" FAMILY OF LIPOLYTIC ENZYMES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X53868; CAA37862.1; -.
DR PIR; A39556; A39556.
DR InterPro; IPR002168; Lipolytic enzyme.
DR InterPro; IPR000379; Ser_estr_site.
DR PROSITE; PS01173; LIPASE_GDXG_HIS; 1.
DR PROSITE; PS01174; LIPASE_GDXG_SER; 1.
KW Hydrolase; Lipid degradation.
FT ACT_SITE 165 165 POTENTIAL.
FT ACT_SITE 239 239 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT MUTAGEN 165 165 H->Q: LOSS OF LIPASE & ESTERASE ACTIVITY.
FT MUTAGEN 239 239 S->A: LOSS OF LIPASE & ESTERASE ACTIVITY.
SQ SEQUENCE 433 AA; 47224 MW; D4E53252BDF12ACD CRC64;

Query Match 80.5%; Score 33; DB 1; Length 433;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
DB 180 FCHTVC 185

RESULT 6
TBG_ENTHI
ID_TBGENTHI STANDARD; PRT; 451 AA.
AC P54301;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tubulin gamma chain (gamma tubulin).
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HM-1; IMSS;
RX MEDLINE=98156770; PubMed=9497056;
RA Ray S.S., Gangopadhyay S.S., Pande G., Samuelson J.C., Boria A.;
RT "Primary structure of Entamoeba histolytica gamma-tubulin and
localisation of amoebic microtubule organising centres.";
RL Mol. Biochem. Parasitol. 90:331-336(1997).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES.
GAMMA TUBULIN IS FOUND AT MICROTUBULE ORGANIZING CENTERS (MTOC)
SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U20322; AAC08441.1; -.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_FtsZ.
DR Pfam; PF00091; tubulin; 1.
DR PRINTS; PR01161; TUBULIN.
DR PROSITE; PS00227; TUBULIN; FALSE_NEG.

```

```
KW Microtubules; GTP-binding.
FT NP_BIND 139 145 GTP (POTENTIAL).
SQ SEQUENCE 451 AA; 51955 MW; 1DAS422E0134B256 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 451;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6
   :||:|
Db 134 FCHSIC 139

RESULT 7
VNC5 PAVL3 STANDARD; PRT; 668 AA.
ID VNC5 PAVL3 STANDARD; PRT; 668 AA.
AC P36311;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Difford N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus LuIII and
RT localization of a unique sequence possibly responsible for its
RT encapsidation pattern.";
RL Virology 192:339-345(1993).
CC -!- FUNCTION: SEEMS NECESSARY FOR VIRAL DNA REPLICATION.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES NONCAPSID PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81888; -; NOT ANNOTATED_CDS.
DR EMBL; M12032; AAAG9567.1; -;
DR PIR; A23008; UYPVIM.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
DR Nonstructural protein; Noncapsid protein; DNA replication;
KW ATP-binding.
KW NP_BIND 399 406 ATP (POTENTIAL).
FT CONFLICT 597 597 I -> L (IN REF. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 672;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
   :||:|
Db 378 CHAIC 382

RESULT 9
VNC5 MUMIV STANDARD; PRT; 672 AA.
ID VNC5 MUMIV STANDARD; PRT; 672 AA.
AC P03134;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NS1.
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
RT parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
CC -!- FUNCTION: SEEMS NECESSARY FOR VIRAL DNA REPLICATION.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES NONCAPSID PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: J02275; AAA67109.1; -
 CC ENBL; V01115; CAA24309.1; ALT INIT.
 CC FIR; A03696; UYPMV.
 CC DR TRANSFAC; T02375; -
 CC DR InterPro; IPR001257; Parvo NS1.
 CC DR Pfam; PF01057; Parvo NS1; 1.
 CC DR Nonstructural protein; Noncapsid protein; DNA replication;
 CC ATP-binding.

FT NP BIND 399 406 ATP (POTENTIAL).
 SQ SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;
 Query Match 80.5%; Score 33; DB 1; Length 672;
 Best Local Similarity 80.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 CHAUC 6
 |||||
 Db 378 CHAIC 382

RESULT 10
 VNCS PAVHH
 ID_VNCS_PAVHH STANDARD; PRT; 672 AA.

AC P03133;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
 GN NS1.

OS Hamster parvovirus H1.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OC NCBI_TaxID=10799;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83112183; PubMed=6823009;
 RA Rhode S.L. III, Paradiso P.R.;

RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
 RT genes by hybrid-arrested translation.";
 RL J. Virol. 45:173-184(1983).

CC -1- FUNCTION: SEEMS NECESSARY FOR VIRAL DNA REPLICATION.
 CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES NONCAPSID PROTEIN FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: X01457; CAA25689.1; -
 CC FIR; A03695; UYPMV1.
 CC DR InterPro; IPR001257; Parvo NS1.
 CC DR Pfam; PF01057; Parvo NS1; 1.
 CC DR Nonstructural protein; Noncapsid protein; DNA replication;
 CC ATP-binding.

KW Nonstructural protein; Noncapsid protein; DNA replication;
 KW ATP-binding.
 FT NP BIND 399 406 ATP (POTENTIAL).
 SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 80.5%; Score 33; DB 1; Length 672;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 CHAUC 6
 |||||
 Db 378 CHAIC 382

RESULT 11

DMSB ECOLI
 ID DMSB_ECOLI STANDARD; PRT; 204 AA.
 AC P18776; P77745;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Anaerobic dimethyl sulfoxide reductase chain B (DMSO reductase iron-
 DE sulfur subunit).
 GN DMSB OR B0895.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-9.
 RX STRAIN=K12 / C600;
 RX MEDLINE=89096500; PubMed=3062312;
 RA Bilous P.T., Cole S.T., Anderson W.F., Weiner J.H.;
 RT "Nucleotide sequence of the dmsABC operon encoding the anaerobic
 RT dimethylsulphoxide reductase of Escherichia coli.";
 RL Mol. Microbiol. 2:785-795(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 RN [4]
 RP EPR SPECTROSCOPY OF IRON-SULFUR CLUSTERS.
 RX MEDLINE=91070067; PubMed=2174699;
 RA Cammack R., Weiner J.H.;
 RT "Electron paramagnetic resonance spectroscopic characterization of
 RT dimethyl sulfoxide reductase of Escherichia coli.";
 RL Biochemistry 29:8410-8416(1990).
 RN [5]
 RP MUTAGENESIS.
 RX MEDLINE=91355180; PubMed=1653010;
 RA Rothery R.A., Weiner J.H.;
 RT "Alteration of the iron-sulfur cluster composition of Escherichia
 RT coli dimethyl sulfoxide reductase by site-directed mutagenesis.";
 RL Biochemistry 30:8296-8305(1991).
 CC -1- FUNCTION: Electron transfer subunit of the terminal reductase
 CC during anaerobic growth on various sulfoxide and N-oxide
 CC compounds.
 CC -1- COFACTOR: BINDS 4 4FE-4S CLUSTERS.
 CC -1- SUBUNIT: THE COMPLEX CONSIST OF THREE SUBUNITS: DMSA, THE
 CC REDUCTASE; DMSB, AN ELECTRON TRANSFER PROTEIN, AND DMSC, A
 CC MEMBRANE ANCHOR PROTEIN.
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

DR ENBL; J03412; AAA83844.1; --

DR ENBL; AE000191; AAG73981.1; --

DR ENBL; D90727; BAA35627.1; --

DR PIR; S03786; S03786.

DR HSP; P00195; ICLF.

DR EcoGene; EG10233; dmsB.

DR InterPro; IPR001450; 4Fe4S_ferredoxin.

DR Pfam; PF00037; fer4; 1.

DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.

KM Electron transport; 4Fe-4S; Iron-sulfur; Complete proteome.

FT INIT MET 0

FT METAL 13 13 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 16 16 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 19 19 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 23 23 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 66 66 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 69 69 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 74 74 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 78 78 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 98 98 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 101 101 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 104 104 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 108 108 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 125 125 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT METAL 128 128 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT METAL 140 140 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT METAL 144 144 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT MUTAGEN 101 101 C-F, S, W, Y; LOSS OF ELECTRON TRANSFER FROM MENAQUINOL TO DMSO.

FT CONFLICT 169 169 P -> PRA (IN REF. 1).

FT SEQUENCE 204 AA; 22738 MW; 90425D307DEF2130 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 204;

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YCHAVC 6

Db 103 YCHMAC 108

RESULT 12

YNFG ECOLI STANDARD; PRT; 204 AA.

AC P77313.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable anaerobic dimethyl sulfoxide reductase chain ynfG (DMSO reductase iron-sulfur subunit ynfG).

GN YNF OR B1589.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OC NCBI TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474 (1997).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=K12;

RC MEDLINE=97251357; PubMed=9097039;

RA Aiba H., Baba T., Fujita K., Hayaishi K., Inada T., Isono K., Itoh T.,

RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y., Nishio Y., Oshino T., Saito N., Sampaio G., Seki Y., Sivasubramanian S., Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;

RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map.";

RL DNA Res. 3:363-377 (1996).

CC -|- FUNCTION: Electron transfer subunit of the terminal reductase during anaerobic growth on various sulfoxide and N-oxide compounds.

CC -|- COFACTOR: BINDS 4 4FE-4S CLUSTERS (By similarity).

CC -|- SUBUNIT: THE COMPLEX CONSIST OF THREE SUBUNITS: YNFF, THE REDUCTASE; YNFG, AN ELECTRON TRANSFER PROTEIN, AND YNFB, A MEMBRANE ANCHOR PROTEIN (Potential).

CC -|- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

CC ENBL; AE000254; AAC74661.1; --

CC ENBL; D90801; BAA15313.1; --

CC ENBL; D90802; BAA15323.1; --

CC HSP; P00195; ICLF.

CC EcoGene; EG13845; ynfG.

DR InterPro; IPR001450; 4Fe4S_ferredoxin.

DR Pfam; PF00037; fer4; 2.

DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.

DR Hypothetical protein; Electron transport; 4Fe-4S; Iron-sulfur; Complete proteome.

KW INIT MET 0 0 BY SIMILARITY.

FT METAL 13 13 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 16 16 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 19 19 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 23 23 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).

FT METAL 66 66 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 69 69 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 74 74 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 78 78 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 98 98 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 101 101 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 104 104 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 108 108 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).

FT METAL 125 125 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT METAL 128 128 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT METAL 140 140 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT METAL 144 144 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).

FT SEQUENCE 204 AA; 22620 MW; 5D3B853AD0E2FDEA CRC64;

Query Match 78.0%; Score 32; DB 1; Length 204;

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YCHAVC 6

Db 103 YCHMAC 108

RESULT 13

DMSB HAEIN STANDARD; PRT; 205 AA.

ID P45003; Q48049;

AC 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Anaerobic dimethyl sulfoxide reductase chain B (DMSO reductase iron-sulfur subunit).

GN DMSB OR H11046.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus
 OX NCBI_TaxID=727;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavang A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geochagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Egan / Serotype B;
 RX MEDLINE=96186920; PubMed=8635740;
 RA Loomore S.M., Shortreed J.M., Coleman D.C., England D.M., Klein M.H.;
 RT "Sequences of the genes encoding the A, B and C subunits of the
 RT Haemophilus influenzae dimethylsulfoxide reductase complex.";
 RL Gene 169:137-138(1996).
 CC -1- FUNCTION: Electron transfer subunit of the terminal reductase
 CC during anaerobic growth on various sulfoxide and N-oxide
 CC compounds (By similarity).
 CC -1- COFACTOR: BINDS 4 4FE-4S CLUSTERS (By similarity).
 CC -1- SUBUNIT: THE COMPLEX CONSIST OF THREE SUBUNITS: DMSA, THE
 CC REDUCTASE; DMSB, AN ELECTRON TRANSFER PROTEIN, AND DMSB, A
 CC MEMBRANE ANCHOR PROTEIN (By similarity).
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U32785; AAC22705.1; -;
 DR EMBL; U26665; AAB06234.1; -;
 DR HSP; P00198; 2FDN.
 DR TIGR; H11046; -;
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.
 DR Pfam; PF00037; fer4; 2.
 DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
 KW Electron transport; 4Fe-4S; Iron-sulfur; Complete proteome.
 FT METAL 13 13 IRON-SULFUR 1 (4FE-4S) (By similarity).
 FT METAL 16 16 IRON-SULFUR 1 (4FE-4S) (By similarity).
 FT METAL 19 19 IRON-SULFUR 1 (4FE-4S) (By similarity).
 FT METAL 23 23 IRON-SULFUR 1 (4FE-4S) (By similarity).
 FT METAL 67 67 IRON-SULFUR 2 (4FE-4S) (By similarity).
 FT METAL 70 70 IRON-SULFUR 2 (4FE-4S) (By similarity).
 FT METAL 75 75 IRON-SULFUR 2 (4FE-4S) (By similarity).
 FT METAL 79 79 IRON-SULFUR 2 (4FE-4S) (By similarity).
 FT METAL 99 99 IRON-SULFUR 3 (4FE-4S) (By similarity).
 FT METAL 102 102 IRON-SULFUR 3 (4FE-4S) (By similarity).
 FT METAL 105 105 IRON-SULFUR 3 (4FE-4S) (By similarity).
 FT METAL 109 109 IRON-SULFUR 3 (4FE-4S) (By similarity).
 FT METAL 126 126 IRON-SULFUR 4 (4FE-4S) (By similarity).
 FT METAL 129 129 IRON-SULFUR 4 (4FE-4S) (By similarity).
 FT METAL 141 141 IRON-SULFUR 4 (4FE-4S) (By similarity).
 FT METAL 145 145 IRON-SULFUR 4 (4FE-4S) (By similarity).
 FT VARIANT 60 60 A -> G (IN STRAIN EGAN).
 SQ SEQUENCE 205 AA; 22927 MW; 006A89D0802BF685 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 205;
 Best Local Similarity 66.7%; Pred. No. 57;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 DB 104 YCHMAC 109
 RESULT 14
 GB01 BOVIN
 ID GB01 BOVIN STANDARD; PRT; 353 AA.
 AC P08239;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Guanine nucleotide-binding protein G(O), alpha subunit 1 (GTP-binding
 DE protein G39).
 GN GNAO1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87204086; PubMed=3106961;
 RA van Meurs K.P., Angus C.W., Lavu S., Kung H.-F., Czarnecki S.K.,
 RA Moss J., Vaughan M.;
 RT "Deduced amino acid sequence of bovine retinal G_o alpha: similarities
 RT to other guanine nucleotide-binding proteins";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:3107-3111(1987).
 RN [2]
 RP SEQUENCE OF 145-205 FROM N.A.
 RX MEDLINE=86287343; PubMed=3090546;
 RA Angus C.W., van Meurs K.P., Tsai S.-C., Adamik R., Miedel M.C.,
 RA Pan Y.-C.E., Kung H.-F., Moss J., Vaughan M.;
 RT "Identification of the probable site of cholesterol-catalyzed ADP-
 RT ribosylation in a G_o alpha-like protein based on cDNA sequence.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5813-5816(1986).
 RN [3]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Cerebellum;
 RA Ovchinnikov Y.A., Slepak V.Z., Pronin A.N., Shlenskii A.B.,
 RA Levina N.B., Voeikov V.L., Bystrov N.S., Severtsova I.V., Lipkin V.M.;
 RT "G39, GTP-binding protein from bovine cerebellum. Amino acid sequence
 RT and nucleotide sequence of the corresponding cDNA.";
 RL Dokl. Biochem. 297:404-406(1987).
 RN [4]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Cerebellum;
 RX MEDLINE=88083643; PubMed=3121394;
 RA Ovchinnikov Y.A., Slepak V.Z., Pronin A.N., Shlenskii A.B.,
 RA Levina N.B., Voeikov V.L., Lipkin V.M.;
 RT "Primary structure of bovine cerebellum GTP-binding protein G39 and
 RT its effect on the adenylate cyclase system.";
 RL FEBS Lett. 226:191-95(1987).
 RN [5]
 RP MYRISTOYLATION.
 RX MEDLINE=87298564; PubMed=3113429;
 RA Schultz A.M., Tsai S.C., Kung H.F., Oroszlan S., Moss J., Vaughan M.;
 RT "Hydroxylamine-stable covalent linkage of myristic acid in G_o alpha,
 RT a guanine nucleotide-binding protein of bovine brain.";
 RL Biochem. Biophys. Res. Commun. 146:1234-1239(1987).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -1- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
 CC (G1/O/T/2).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M16116; AAA30530.1; --
EMBL; M14489; AAA30551.1; --
EMBL; X12924; CAA31391.1; --
PIR; S00213; RGH001.
HSSP; P10824; IBOF.
InterPro; IPR001019; Gprotein.alpha.
InterPro; IPR001230; Prenyl_site.
Pfam; PF00503; G-alpha; 1.
PRINTS; PR000281; GPROTEIN.
ProDom; PD000281; Gprotein.alpha; 1.
SMART; SM00275; G-alpha; 1.
GTP-binding; Transducer; Multigene family; Alternative splicing;
ADP-ribosylation; Myristate; Palmitate; Lipoprotein.
FT INIT MET 0 0 MYRISTATE.
FT LIPID 1 1 PALMITATE. (BY SIMILARITY).
FT NP_BIND 39 46 GTP (BY SIMILARITY).
FT NP_BIND 200 204 GTP (BY SIMILARITY).
FT MOD_RES 269 272 GTP (BY SIMILARITY).
FT MOD_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX).
FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
FT CONFLICT 252 252 S -> F (IN REF. 3).
FT CONFLICT 262 262 I -> T (IN REF. 3 AND 4).
SQ SEQUENCE 353 AA; 39947 MW; 3E0FB2B947C37FE0 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 353;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YCHAVC 6
|||
Db 319 YCHMTC 324

RESULT 15
GB01 HUMAN STANDARD; PRT; 353 AA.
AC P09471.
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein G(i), alpha subunit 1.
GN GNAO1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91195273; PubMed=1901650;
RA Tsukamoto T., Toyama R., Itoh H., Kozasa T., Matsuo M., Kaziro Y.;
RT "Structure of the human gene and two rat cDNAs encoding the alpha
chain of GTP-binding regulatory protein Go: two different mRNAs are
generated by alternative splicing."
RL Proc. Natl. Acad. Sci. U.S.A. 88:2974-2978(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88134191; PubMed=3124840;
RA Lavu S., Clark J., Swarup R., Matshushima K., Paturu K., Moss J.,
RA Kung H.-F.;
RT "Molecular cloning and DNA sequence analysis of the human guanine
nucleotide-binding protein Go alpha."
RL Biochem. Biophys. Res. Commun. 150:811-815(1988).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE

CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
CC -1- SUBUNIT: G-PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-1 (SHOWN HERE) AND ALPHA-2
CC (AC P29777); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
CC (G1/O7/Z1).
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M60164; AAA52585.1; --
EMBL; M60156; AAA52585.1; JOINED.
EMBL; M60157; AAA52585.1; JOINED.
EMBL; M60158; AAA52585.1; JOINED.
EMBL; M60159; AAA52585.1; JOINED.
EMBL; M60160; AAA52585.1; JOINED.
EMBL; M60163; AAA52585.1; JOINED.
EMBL; M19182; AAA52584.1; --
EMBL; M19184; AAA52584.1; JOINED.
PIR; A40436; RGHU01.
HSSP; P10824; IBOF.
Pfam; PF00503; G-alpha; 1.
Genew; HGNC:4389; GNAO1.
MIM; 139311; --
InterPro; IPR001019; Gprotein.alpha.
InterPro; IPR001230; Prenyl_site.
Pfam; PF00503; G-alpha; 1.
PRINTS; PR000281; Gprotein.alpha; 1.
ProDom; PD000281; Gprotein.alpha; 1.
SMART; SM00275; G-alpha; 1.
GTP-binding; Transducer; Multigene family; Alternative splicing;
ADP-ribosylation; Myristate; Palmitate; Lipoprotein.
FT INIT MET 0 0 MYRISTATE.
FT LIPID 1 1 PALMITATE. (BY SIMILARITY).
FT NP_BIND 39 46 GTP (BY SIMILARITY).
FT NP_BIND 200 204 GTP (BY SIMILARITY).
FT MOD_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX).
FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
FT CONFLICT 15 15 S -> G (IN REF. 2).
FT CONFLICT 170 170 T -> L (IN REF. 2).
FT CONFLICT 217 217 D -> E (IN REF. 2).
SQ SEQUENCE 353 AA; 39919 MW; 64C6DCDC67A549B0 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 353;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YCHAVC 6
|||
Db 319 YCHMTC 324

RESULT 16
GB01 MOUSE STANDARD; PRT; 353 AA.
AC P18872;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein G(O), alpha subunit 1.
GN GNAO1 OR GNAO OR GNAO.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Spermatid;
 RX MEDLINE=90370808; PubMed=1697681;
 RA Strachmann M., Wilkie T.M., Simon M.I.;
 RT "Alternative splicing produces transcripts encoding two forms of the
 RL alpha subunit of GTP-binding protein Go.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6477-6481(1990).
 CC -|- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -|- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
 CC -|- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-1 (SHOWN HERE) AND ALPHA-2
 CC (AC P18873); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -|- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
 CC (G(I/O/T/Z)).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M36777; AAA37645.1; -;
 CC PIR; A36038; RGM501.
 CC HSSP; P10824; 1BOF.
 CC MGD; MG1:95775; Gnao.
 CC InterPro; IPR001049; Gprotein alpha.
 CC Refam; PF00503; G-alpha; 1.
 CC PRINTS; PR00318; GPROTEINA.
 CC ProDom; PD000281; G-alpha; 1.
 CC SMART; SM00275; G-alpha; 1.
 CC GTP-binding; Transducer; Multigene family; Alternative splicing;
 KW ADP-ribosylation; Myristate; Palmitate; Lipoprotein.
 FT INIT_MET 0 0 MYRISTATE (BY SIMILARITY).
 FT LIPID 1 1 PALMITATE (BY SIMILARITY).
 FT NP_BIND 39 46 GTP (BY SIMILARITY).
 FT NP_BIND 200 204 GTP (BY SIMILARITY).
 FT NP_BIND 269 272 GTP (BY SIMILARITY).
 FT MOD_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX).
 FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
 SQ SEQUENCE 353 AA; 39953 MW; 9CE471097205BBA2 CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 353;
 Best Local Similarity 66.7%; Pred. No. 86;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 |||||
 Db 319 YCHWTC 324
 RESULT 17
 GB01_RAT
 ID GB01_RAT STANDARD; PRT; 353 AA.
 AC P0490;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Guanine nucleotide-binding protein G(O), alpha subunit 1.
 GN GNAO1 OR GNAO OR GNAO
 OS Rattus norvegicus (Rat), and
 OS Cricetus longicaudatus (Long-tailed hamster) (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116, 10030;

RN RP SEQUENCE FROM N.A.
 RC SPECIES=Rat;
 RX MEDLINE=88007678; PubMed=2820999;
 RA Jones D.T., Reed R.R.;
 RT "Molecular cloning of five GTP-binding protein cDNA species from rat
 RL olfactory neuroepithelium.";
 RL J. Biol. Chem. 262:14241-14249(1987).
 RN [2]
 RP SEQUENCE OF 44-353 FROM N.A.
 RC SPECIES=Rat;
 RX MEDLINE=86233317; PubMed=3086867;
 RA Itoh H., Kozasa T., Nagata S., Nakamura S., Katada T., Ui M., Iwai S.,
 RA Ohtsuka E., Kawaaki H., Suzuki K., Kaziro Y.;
 RT "Molecular cloning and sequence determination of cDNAs for alpha
 RL subunits of the guanine nucleotide-binding proteins Gs, Gi, and Go
 RL from rat brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3776-3780(1986).
 RN [3]
 RP SEQUENCE OF 54-66; 113-128 AND 162-175.
 RC SPECIES=Rat;
 RX MEDLINE=90231435; PubMed=2158629;
 RA Strittmatter S.M., Valenzuela D., Kennedy T.E., Neer E.J.,
 RA Fishman M.C.;
 RT "GO is a major growth cone protein subject to regulation by GAP-43.";
 RL Nature 344:836-841(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.longicaudatus;
 RX MEDLINE=90293067; PubMed=2113531;
 RA Hsu W.H., Rudolph U., Sanford J., Bertrand P., Olate J.,
 RA Nelson C., Moss L.G., Boyd A.E. III, Codina J., Birnbaumer L.;
 RT "Molecular cloning of a novel splice variant of the alpha subunit of
 RL the mammalian Go protein.";
 RL J. Biol. Chem. 265:11220-11226(1990).
 RN [5]
 RP PALMITOYLATION.
 RC SPECIES=Rat;
 RX MEDLINE=91249349; PubMed=8484716;
 RA Parenti M., Vigano M.A., Newman C.M.H., Milligan G., Magee A.I.;
 RT "A novel N-terminal motif for palmitoylation of G-protein alpha
 RL subunits.";
 RL Biochem. J. 291:349-353(1993).
 RN [6]
 RP DEMYDICATION OF ASN-345.
 RX MEDLINE=99145549; PubMed=9990023;
 RA Exner T., Jensen O.N., Mann M., Kleus C., Nurnberg B.;
 RT "Posttranslational modification of Galphao1 generates Galphao3, an
 RL abundant G protein in brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1327-1332(1999).
 CC -|- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -|- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
 CC -|- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -|- ALTERNATIVE PRODUCTS: TWO FORMS OF G(O) ALPHA SUBUNITS ARE
 CC PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -|- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
 CC (G(I/O/T/Z)).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M17526; AAA40826.1; -;
 CC EMBL; M12671; AAA41262.1; -;
 CC EMBL; M33661; AAA36987.1; -;
 CC PIR; A42228; RGHYO1.

DR PIR; C40436; RGPTOL.
 DR HSP; P10824; IBOF.
 DR InterPro; IPR001019; G-protein alpha.
 DR InterPro; IPR001230; Prenyl_site.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR ProDom; PD000281; G-protein_alpha; 1.
 DR SMART; SM00275; G-alpha; 1.
 KW GTP-binding; Transducer; Multigene family; Alternative splicing;
 KW ADP-ribosylation; Myristate; Palmitate; Lipoprotein.
 FT INIT MET 0 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
 FT LIPID 2 2 PALMITATE.
 FT NP_BIND 39 46 GTP (BY SIMILARITY).
 FT NP_BIND 200 204 GTP (BY SIMILARITY).
 FT NP_BIND 269 272 GTP (BY SIMILARITY).
 FT MOD_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX).
 FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
 FT MOD_RES 345 345 DEAMIDATION (IN ISOFORM ALPHA-3).
 SQ SEQUENCE 353 AA; 39937 MW; 7CAED10CD4A2BB07 CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 353;
 Best Local Similarity 66.7%; Pred. No. 88;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 DB 319 YCHMTC 324
 RESULT 18
 GB0_CAEEL STANDARD; PRT; 353 AA.
 ID GB0_CAEEL
 AC P51875; Q18205;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Guanine nucleotide-binding protein G(O), alpha subunit.
 GN GOA-1 OR C26C6.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=9132211; PubMed=1907494;
 RA Lochrie M.A., Mendel J.E., Sternberg P.W., Simon M.I.;
 RT "Homologous and unique G protein alpha subunits in the nematode
 Caenorhabditis elegans.";
 RL Cell Regul. 2:135-154(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Cuppen E., Jansen G., Plasterk R.H.A.;
 RT "Interaction analysis of the complete G-alpha subfamily of
 heterotrimeric G proteins from Caenorhabditis elegans.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA McMurray A.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 SIGNALING SYSTEMS.
 CC -1- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC Interacts with gpb-2.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
 (G(I/O/T/2)).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M38251; AAA28059.1;
 DR EMBL; AY008140; AAG32093.1;
 DR EMBL; Z72503; CA96595.1;
 DR WormPep; C26C6.2; CE05311.
 DR HSP; P04896; IAZT.
 DR InterPro; IPR001019; G-protein_alpha.
 DR InterPro; IPR001230; Prenyl_site.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR ProDom; PD000281; G-protein_alpha; 1.
 DR SMART; SM00275; G-alpha; 1.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
 KW Myristate; Palmitate; Lipoprotein.
 FT INIT MET 0 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
 FT LIPID 2 2 PALMITATE (BY SIMILARITY).
 FT NP_BIND 39 46 GTP (BY SIMILARITY).
 FT NP_BIND 200 204 GTP (BY SIMILARITY).
 FT NP_BIND 269 272 GTP (BY SIMILARITY).
 FT MOD_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX).
 FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
 FT CONFLICT 173 173 D -> H (IN REF. 1).
 FT CONFLICT 233 233 V -> L (IN REF. 1).
 SQ SEQUENCE 353 AA; 40320 MW; A746D1A641296F3E CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 353;
 Best Local Similarity 66.7%; Pred. No. 88;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 DB 319 YCHMTC 324
 RESULT 19
 GB0_HELTI STANDARD; PRT; 353 AA.
 ID GB0_HELTI
 AC P51877;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Guanine nucleotide-binding protein G(O), alpha subunit.
 OS Helisoma trivolvis (Snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 OC Planorbidae; Helisoma.
 OX NCBI_TaxID=27815;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon red; TISSUE=Central ganglion;
 RX MEDLINE=94138304; PubMed=750312;
 RA Durgarian S., Bahls F., Richmond J., Doyle R.T., Larson D.D.,
 RA Haydon P.G.;
 RT "Roles for arachidonic acid and GTP-binding proteins in synaptic
 transmission.";
 RL J. Physiol. (Paris) 87:123-137(1993).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 SIGNALING SYSTEMS.
 CC -1- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
 (G(I/O/T/2)).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L18921; AAC41539.1; -
 DR HSSP; P04896; IAZT.
 DR InterPro; IPR001019; Gprotein_alpha.
 DR InterPro; IPR001230; Prenyl_site.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR ProDom; PD000281; Gprotein_alpha; 1.
 DR SMART; SM00275; G-alpha; 1.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
 KW Myristate; Palmitate; Lipoprotein.
 FT INIT MET 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
 FT LIPID 2 2 PALMITATE (BY SIMILARITY).
 FT NP_BIND 39 46 GTP (BY SIMILARITY).
 FT NP_BIND 200 204 GTP (BY SIMILARITY).
 FT NP_BIND 269 272 GTP (BY SIMILARITY).
 FT MOD_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX).
 FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP).
 SQ SEQUENCE 353 AA; 40127 MW; 04FBC7C4A9DCDD7C CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 353;
 Best Local Similarity 66.7%; Pred. No. 88;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 DB 319 YCHQTC 324

RESULT 20
 GB01 LYMTST STANDARD; PRT; 353 AA.
 AC P30683;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Guanine nucleotide-binding protein G(O), alpha subunit.
 OS Lymnaea stagnalis (Great pond snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 OC Lymnaeidae; Lymnaea.
 OC NCBI_TaxID=6523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CNS;
 RX MEDLINE=93106153; PubMed=1468550;
 RA Knol J.C., Weidemann W., Planta R.J., Vreugdenhil E.,
 RA van Heerikhuizen H.;
 RT "Molecular cloning of G protein alpha subunits from the central
 RT nervous system of the mollusc Lymnaea stagnalis";
 RL FEBS Lett. 314:215-219 (1992).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -1- SUBUNIT: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
 CC (G(T/O/T/2)).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; Z15094; CAA78806.1; -
 DR PIR; S25589; S25589.
 DR PIR; S27014; S27014.
 DR HSSP; P04896; IAZT.
 DR InterPro; IPR001019; Gprotein_alpha.
 DR InterPro; IPR001230; Prenyl_site.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR ProDom; PD000281; Gprotein_alpha; 1.
 DR SMART; SM00275; G-alpha; 1.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
 KW Myristate; Palmitate; Lipoprotein.
 FT INIT MET 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
 FT LIPID 2 2 PALMITATE (BY SIMILARITY).
 FT NP_BIND 39 46 GTP (BY SIMILARITY).
 FT NP_BIND 200 204 GTP (BY SIMILARITY).
 FT NP_BIND 269 272 GTP (BY SIMILARITY).
 FT MOD_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT MOD_RES 350 350 ADP-RIBOSYL[1] (BY ACTION OF IAP)
 SQ SEQUENCE 353 AA; 40140 MW; 01925758C05F654E CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 353;
 Best Local Similarity 66.7%; Pred. No. 88;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YCHAVC 6
 DB 319 YCHQTC 324

RESULT 21
 GB01 DROME STANDARD; PRT; 354 AA.
 AC P16378; P16707; Q9V5L5; Q9V5L6;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Guanine nucleotide-binding protein G(O), alpha subunit 47A.
 GN G-O-ALPHA-47A OR G-OA47A OR CG2204.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.
 RX MEDLINE=90036957; PubMed=2509464;
 RA Thambi N.C., Quan F., Wolfgang W.J., Spiegel A., Forte M.;
 RT "Immunological and molecular characterization of G alpha-like
 RT proteins in the Drosophila central nervous system";
 RL J. Biol. Chem. 264:18552-18560 (1989).
 RN [2]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.
 RC TISSUE=Head;
 RX MEDLINE=90036955; PubMed=2509462;
 RA Yoon J., Shortridge R.D., Bloomquist B.T., Schneuwly S., Perdev M.H.,
 RA Pak W.L.;
 RT "Molecular characterization of Drosophila gene encoding G alpha
 RT subunit homolog";
 RL J. Biol. Chem. 264:18536-18543 (1989).
 RN [3]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=90036956; PubMed=2509463;
 RA de Sousa S.M., Hoveland L.L., Yarfitz S., Hurley J.B.;
 RT "The Drosophila G alpha-like G protein gene produces multiple
 RT transcripts and is expressed in the nervous system and in ovaries";
 RL J. Biol. Chem. 264:18544-18551 (1989).
 RN [4]

SEQUENCE FROM N.A. (ISOFORM G-OA47A-2), AND TISSUE SPECIFICITY.
RX MEDLINE=92096479; PubMed=2519611;
RA Schmidt C.J., Garén-Fazio S., Chow Y.K., Neer E.J.;
RT "Neuronal expression of a newly identified Drosophila melanogaster G
protein alpha 0 subunit";
PL Cell Regul. 1:125-134(1989).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM G-OA47A-1 AND G-OA4A-2).
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.C., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Baou A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Buam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laiko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
CC -|- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -|- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
CC -|- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; G-OA47A-1 (SHOWN HERE) AND G-
CC OA47A-2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -|- TISSUE SPECIFICITY: EXPRESSED PRIMARILY IN NEURONAL CELL BODIES IN
CC THE BRAIN, OPTIC LOBE, AND THORACIC AND ABDOMINAL GANGLIA. ALSO
CC EXPRESSED IN ANTENNA, OOCYTES AND OVARIAN NERVE CELLS.
CC -|- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT, HIGHEST
CC LEVEL IN ADULTS.
CC -|- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
CC (G(I/O/T/Z)).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M29731; AAA28580.1; ..

DR EMBL; M29732; AAA28581.1; ..
DR EMBL; M31203; AAA28586.1; ..
DR EMBL; M29601; AAA28586.1; JOINED.
DR EMBL; M31198; AAA28586.1; JOINED.
DR EMBL; M31199; AAA28586.1; JOINED.
DR EMBL; M31200; AAA28586.1; JOINED.
DR EMBL; M31201; AAA28586.1; JOINED.
DR EMBL; M31202; AAA28586.1; JOINED.
DR EMBL; M29602; AAA28587.1; ..
DR EMBL; M30151; AAA28587.1; ..
DR EMBL; M30152; AAA28585.1; ..
DR EMBL; M31129; AAA28583.1; ..
DR EMBL; M86660; AAA28577.1; ..
DR EMBL; AS0003829; AAF58789.1; ..
DR EMBL; AS0003829; AAF58789.1; ..
DR PIR; A34304; RGFF02.
DR PIR; B34304; RGFF02.
DR HSSP; P04896; IAZT.
DR FlyBase; FBgn0001122; G-O-alpha-47A.
DR InterPro; IPR001019; Gprotein_alpha.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR003575; Small_GTPase.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEIN.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G-alpha; 1.
DR SMART; SM00010; small_GTPase; 1.
DR GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW Alternative splicing.
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 201 205 GTP (BY SIMILARITY).
FT NP_BIND 270 273 GTP (BY SIMILARITY).
FT MOD_RES 179 179 ADP-RIBOSYL[1] (BY ACTION OF CTX).
FT MOD_RES 351 351 ADP-RIBOSYL[1] (BY ACTION OF IAP).
FT VARSPLIC 4 21 AQAEEERAAAEERLIER -> TTSAEERAAIORSKOIEK
FT CONFLICT 88 88 M -> I (IN REF. 3; AAA28585).
FT SEQUENCE 354 AA; 40476 MW; 3C5DA142B4CF7DD2 CRC64;
Query Match 78.0%; Score 32; DB 1; Length 354;
Best Local Similarity 76.7%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 YCHAVC 6
DB 320 YCHMTC 325
RESULT 22
GB0_LOCM1 STANDARD; PRT; 354 AA.
ID GB0_LOCM1
AC P38404;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein G(O), alpha subunit.
OS Locusta migratoria [Migratory locust].
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90380780; PubMed=2119206;
RA Raming K., Krieger J., Breer H.;
RT "Molecular cloning, sequencing and expression of cDNA encoding a G0-
protein from insect";
RL Cell. Signal. 2:311-321(1990).
CC -|- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -|- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
CC -|- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).

CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
(G(I/O/T/Z)).
CC PIR; A61035; A61035.
DR HSP; P04896; LAZT.
DR InterPro; IPR001019; Gprotein_alpha.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR003575; Small_GTPase.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR ProDom; PD000381; Gprotein_alpha; 1.
DR SMART; SM00275; G-alpha; 1.
DR SMART; SM00010; small_GTPase; 1.
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 201 205 GTP (BY SIMILARITY).
FT NP_BIND 270 273 GTP (BY SIMILARITY).
FT MOD_RES 179 179 ADP-RIBOSYL[1] (BY ACTION OF CTX).
FT MOD_RES 351 351 ADP-RIBOSYL[1] (BY ACTION OF IAP).
SQ SEQUENCE 354 AA; 40337 MW; 9F46951E95D8EEAE CRC64;

Query Match 78.0%; Score 32; DB 1; Length 354;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCHAVC 6
DB 320 YCHWTC 325

RESULT 23
GB0 PATYE
ID GB0 PATYE STANDARD; PRT; 356 AA.
AC O15576;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein G(O), alpha subunit.
GN SCGOA.
OS Patinopecten yessoensis (Ezo giant scallop) (Yesso scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
OC Pectinoidea; Pectinidae; Mizuhopecten.
OX NCBI_TaxID=6573;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=57435252; PubMed=9287291;
RA Kojima D., Terakita A., Ishikawa T., Tsukahara Y., Maeda A.,
RA Shichida Y.;
RT "A novel G α mediated phototransduction cascade in scallop visual
RT cells";
RL J. Biol. Chem. 272:22979-22989(1997).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- FUNCTION: THE G(O) PROTEIN FUNCTION IS NOT CLEAR.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 1
(G(I/O/T/Z)).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB006457; BRA22220.1; --
DR HSP; P04896; LAZT.
DR InterPro; IPR001019; Gprotein_alpha.
DR InterPro; IPR001230; Prenyl_site.

DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G-alpha; 1.
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW Myristate; Palmitate; Lipoprotein.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT LIPID 2 2 PALMITATE (BY SIMILARITY).
FT NP_BIND 39 46 GTP (BY SIMILARITY).
FT NP_BIND 203 207 GTP (BY SIMILARITY).
FT NP_BIND 272 275 GTP (BY SIMILARITY).
FT MOD_RES 181 181 ADP-RIBOSYL[1] (BY SIMILARITY).
FT MOD_RES 353 353 ADP-RIBOSYL[1] (BY SIMILARITY).
SQ SEQUENCE 356 AA; 40623 MW; C668F5E5BDEE376E CRC64;

Query Match 78.0%; Score 32; DB 1; Length 356;
Best Local Similarity 66.7%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCHAVC 6
DB 322 YCHQTC 327

RESULT 24
TOH2 CAEEL
ID TOH2 CAEEL STANDARD; PRT; 592 AA.
AC P98060;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc metalloproteinase toh-2 precursor (EC 3.4.24.-).
GN TOH-2 OR R151.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulton L.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U00036; AAK29859.1; --
DR HSP; P28825; IIAF.
DR MEROPS; M12.0PA;
DR WormPep; R151.5; CE27200.
DR InterPro; IPR001506; Astacin.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000130; Zn_MTPpeptidase.
DR Pfam; PF00090; tep_1; 1.
DR Pfam; PF0431; CUB; 1.
DR Pfam; PF01400; Astacin; 1.
DR PRINTS; PR00480; ASTACIN.

```
DR SMART; SM00042; CUB; 1.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS00142; ZINC-PROTEASE; 1.
KW Hydrolase; Metalloprotease; EGF-like domain; Zinc; signal.
FT SIGNAL 1 18
FT CHAIN 19 592
FT DOMAIN 321 361
FT DOMAIN 371 487
FT DOMAIN 490 541
FT METAL 222 223
FT ACT SITE 223 223
FT METAL 226 226
FT METAL 232 232
FT CARBOHYD 167 167
FT CARBOHYD 438 438
SQ SEQUENCE 592 AA; 67257 MW; 01F2633B0BE53443 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 592;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAV 6
DB 324 YCSAVC 329

RESULT 25
SZ1B BRARE STANDARD; PRT; 778 AA.
AC Q3W686;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Semaphorin 1B precursor (Semaphorin 1B) (Sema-21B).
GN SEMA21B OR SEMA3AB.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9425174; PubMed=10495275;
RA Roos M., Schachner M., Bernhardt R.R.;
RT "Zebrafish semaphorin 1b inhibits growing motor axons in vivo.";
RL Mech. Dev. 87:103-117(1999).
CC -1- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
CC VENTRALLY EXTENDING MOTOR AXONS.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF083382; AAD28103.1;
CC ZFIN; ZDB-GENE-991209-6; semajab.
DR
```

```
DR InterPro; IPR003599; Ig
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01437; PSI; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
Developmental protein; Glycoprotein.
FT SIGNAL 1 17
FT CHAIN 18 778
FT DOMAIN 241 539
FT DOMAIN 645 723
FT DOMAIN 721 776
FT DISULFID 652 716
FT CARBOHYD 54 54
FT CARBOHYD 127 127
FT CARBOHYD 593 593
SQ SEQUENCE 778 AA; 88904 MW; 4D36F4323AE21895 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 778;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAV 5
DB 651 YCHAV 655
```

Search completed: June 20, 2003, 20:52:51
Job time : 9.22727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:42:49 ; Search time 31.6364 Seconds
(without alignments)
39.078 Million cell updates/sec

Title: US-10-105-008-10
Perfect score: 41
Sequence: 1 YHAYC 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database :

SPTREMBL 21:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	90.2	400	2 Q9L514	Q9L514 psychrobact
2	36	87.8	187	16 Q8X6M2	Q8X6M2 escherichia
3	36	87.8	428	10 Q9M1Q8	Q9M1Q8 arabidopsis
4	36	87.8	514	11 Q8R0D0	Q8R0D0 mus musculus
5	36	87.8	519	4 Q9Y5Q8	Q9Y5Q8 homo sapien
6	36	87.8	519	4 Q9H4P2	Q9H4P2 homo sapien
7	36	87.8	520	11 Q8R2T8	Q8R2T8 mus musculus
8	35	85.4	795	4 Q96PV6	Q96PV6 homo sapien
9	34	82.9	114	11 Q9D974	Q9D974 mus musculus
10	34	82.9	170	12 Q910I3	Q910I3 human cytom
11	34	82.9	235	10 Q8SEAL	Q8SEAL oryza sativ
12	34	82.9	238	10 Q9FPEL	Q9FPEL arabidopsis
13	34	82.9	244	10 Q9SHD7	Q9SHD7 arabidopsis
14	34	82.9	380	5 Q9VH34	Q9VH34 drosophila
15	34	82.9	381	5 Q8T9I3	Q8T9I3 drosophila
16	34	82.9	424	10 Q9FRE7	Q9FRE7 oryza sativ

ALIGNMENTS

RESULT 1
ID Q9L514 PRELIMINARY; PRT; 400 AA.
AC Q9L514; (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
GN LIPase.

17	34	82.9	537	6	Q9MZU5	Q9mzu5 sus acrofa
18	34	82.9	646	17	Q9GZZ0	Q9gzz0 sulfolobus
19	34	82.9	1045	5	Q8T3A6	Q8t3a6 caenorhabdi
20	34	82.9	1070	5	Q8T3A7	Q8t3a7 caenorhabdi
21	34	82.9	1111	5	Q9XWD6	Q9xwd6 caenorhabdi
22	34	82.9	3107	12	P87587	P87587 citrus tris
23	34	82.9	3115	12	O10467	O10467 citrus tris
24	34	82.9	3115	12	Q9IFX0	Q9ifx0 citrus tris
25	34	82.9	3115	12	Q9WID7	Q9wid7 citrus tris
26	34	82.9	3132	12	Q9DTG5	Q9dtg5 citrus tris
27	34	82.9	5636	5	Q9N9M2	Q9n9m2 leishmania
28	33	80.5	273	16	Q8Y2B0	Q8y2b0 ralestonia s
29	33	80.5	352	13	Q9DED3	Q9ded3 oncorhynch
30	33	80.5	374	3	Q03289	Q03289 saccharomyc
31	33	80.5	379	17	Q8TK11	Q8tk11 methanosarc
32	33	80.5	383	5	O76856	O76856 dictyostell
33	33	80.5	397	12	Q993M6	Q993m6 autonomus
34	33	80.5	402	12	Q9PY95	Q9py95 human rotav
35	33	80.5	413	5	Q23015	Q23015 caenorhabdi
36	33	80.5	421	16	Q8YUQ5	Q8yug5 anabaena sp
37	33	80.5	573	10	Q9C735	Q9c735 arabidopsis
38	33	80.5	575	10	Q9LYU4	Q9lyu4 arabidopsis
39	33	80.5	578	10	Q49301	Q49301 arabidopsis
40	33	80.5	581	10	Q9FHN7	Q9fhn7 arabidopsis
41	33	80.5	599	10	Q948Q3	Q948q3 physcomitre
42	33	80.5	665	12	O71159	O71159 kilham rat
43	33	80.5	672	12	P88899	P88899 kilham rat
44	33	80.5	672	12	Q83429	Q83429 mouse parvo
45	33	80.5	721	12	Q84365	Q84365 murine minu
46	33	80.5	721	12	Q84363	Q84363 murine minu
47	33	80.5	969	5	Q8T3T3	Q8t3t3 toxoplasma
48	32	78.0	82	12	Q94461	Q94461 paramecium
49	32	78.0	109	11	Q8R4R5	Q8r4r5 rattus norv
50	32	78.0	111	10	Q8S0R8	Q8s0r8 oryza sativ
51	32	78.0	113	2	Q9S124	Q9s124 escherichia
52	32	78.0	157	12	Q98434	Q98434 paramecium
53	32	78.0	205	16	Q82PJ8	Q82pj8 salmonella
54	32	78.0	205	16	Q8Z6V7	Q8z6v7 salmonella
55	32	78.0	205	16	Q8XFB2	Q8xfb2 salmonella
56	32	78.0	205	16	Q9X6B5	Q9x6b5 yersinia pe
57	32	78.0	205	16	Q8X4Q0	Q8x4q0 escherichia
58	32	78.0	205	16	Q8X4K5	Q8x4k5 escherichia
59	32	78.0	206	16	Q9CK78	Q9ck78 pasteurella
60	32	78.0	248	16	Q92MJ7	Q92mj7 rhizobium m
61	32	78.0	354	5	Q9ML93	Q9ml93 octopus vul
62	32	78.0	354	13	Q93564	Q93564 xenopus lae
63	32	78.0	354	13	O13011	O13011 xenopus lae
64	32	78.0	440	10	Q9M6E2	Q9m6e2 taxus cuspi
65	32	78.0	440	10	Q8W169	Q8w169 taxus bacca
66	32	78.0	466	16	Q83526	Q83526 treponema p
67	32	78.0	502	5	Q9U2G6	Q9u2g6 caenorhabdi
68	32	78.0	530	10	Q9LQ87	Q9lqb7 arabidopsis
69	32	78.0	555	13	Q90354	Q90354 cyprinus ca
70	32	78.0	590	5	Q19140	Q19140 caenorhabdi
71	32	78.0	733	10	Q94B71	Q94b71 arabidopsis
72	32	78.0	760	10	Q9LQV7	Q9lqv7 arabidopsis
73	31	75.6	71	2	Q9XB25	Q9xb25 klebsiella
74	31	75.6	71	11	Q60495	Q60495 cavia sp. a
75	31	75.6	83	13	P79707	P79707 cynops pyr

OS Psychrobacter sp. Stl.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Psychrobacter.
 OX NCBI_TaxID=125076;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ST1;
 RA Kulakova L., Galkin A., Kurihara T., Yoshimura T., Baaki N.;
 RT "Cold-active lipase from psychrophilic bacteria Psychrobacter sp. Stl.
 RT gene cloning and protein characterization."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF260707; AAF70342.1; -;
 DR InterPro; IPR002168; Lipolytic enzyme.
 DR InterPro; IPR000379; Serine site.
 DR InterPro; IPR000130; Zn_MiPeptidase.
 DR PROSITE; PS01173; LIPASE GDHG HIS; 1.
 DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN 1.
 SQ SEQUENCE 400 AA; 43683 MW; 108D59146AB6FDB CRC64;

Query Match 90.2%; Score 37; DB 2; Length 400;
 Best Local Similarity 83.3%; Pred. No. 26;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YCHAVC 6
 :|||:
 Db 157 FCHAVC 162

RESULT 2

OQ8X6M2 PRELIMINARY; PRT; 187 AA.
 ID OQ8X6M2
 AC OQ8X6M2;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Prepilin peptidase dependent protein B.
 GN PPB OR Z142 OR EC3562.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimmalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 DR EMBL; AE005511; AAG57936.1; -;
 DR EMBL; AF002563; BAB37105.1; -;
 DR InterPro; IPR001120; Prok N methyltn.
 DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
 KW Complete proteome;
 SQ SEQUENCE 187 AA; 20525 MW; 89AC9E69E8C42D64 CRC64;

Query Match 87.8%; Score 36; DB 16; Length 187;
 Best Local Similarity 66.7%; Pred. No. 22;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 YCHAVC 6
 :|||:
 Db 67 YCHGIC 72

RESULT 3

OQ9M1Q8 PRELIMINARY; PRT; 428 AA.
 ID OQ9M1Q8
 AC OQ9M1Q8;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Hypothetical 47.6 kDa protein.
 GN T17J13.120.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rieger M., Mueller-Auer S., Zipp M., Schaefer M., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL138651; CAB71876.1; -;
 DR InterPro; IPR003480; Transferase.
 DR Pfam; PF02458; Transferase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 428 AA; 47553 MW; 66C5155A161B26CA CRC64;

Query Match 87.8%; Score 36; DB 10; Length 428;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YCHAVC 6
 :|||:
 Db 152 FCHAVC 157

RESULT 4

OQ8R0D0 PRELIMINARY; PRT; 514 AA.
 ID OQ8R0D0
 AC OQ8R0D0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similar to general transcription factor IIIC, polypeptide 5
 DE (63kD).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=EYE;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC027055; AAH27055.1; -;
 SQ SEQUENCE 514 AA; 59868 MW; 39806677F99F4765 CRC64;

Query Match 87.8%; Score 36; DB 11; Length 514;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YCHAVC 6
 :|||:
 Db 70 YCHPVC 75

RESULT 5

Q9Y5Q8 PRELIMINARY; PRT; 519 AA.
 ID Q9Y5Q8
 AC Q9Y5Q8
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Transcription factor IIIIC63.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99303787; PubMed=10373544;
 RA Haieh Y.J., Wang Z., Kovelman R., Roeder R.G.;
 RT "Cloning and characterization of two evolutionarily conserved subunits
 of TFIIC102 and TFIIC63) of human TFIIC and their involvement in
 functional interactions with TFIIB and RNA polymerase III.";
 RL Mol. Cell. Biol. 13:4944-4952(1999).
 DR EMBL; AF133124; AAD41476.1; -;
 SQ SEQUENCE 519 AA; 59629 MW; 0889AB42B30F59D6 CRC64;

Query Match 87.8%; Score 36; DB 4; Length 519;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
 |||||
 DB 70 YCHPVC 75

RESULT 6

Q9H4P2 PRELIMINARY; PRT; 519 AA.
 ID Q9H4P2
 AC Q9H4P2
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Clone CDABP0017 mRNA sequence (Hypothetical 59.6 kDa protein)
 DE (Unknown) (Protein for MGC:18246).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96207227; PubMed=8619474;
 RA Anderson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;
 RT "A 'double adaptor' method for improved shotgun library
 construction.";
 RL Anal. Biochem. 236:107-113(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97264341; PubMed=9110174;
 RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RT "Large-scale concatenation cDNA sequencing.";
 RL Genome Res. 7:353-358(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Zhou J., Yu W., Tang H., Mei G., Tsang Y.T.M., Bouck J., Gibbs R.A.,
 RA Margolin J.F.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKIN;
 RA Strausberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY007123; AGO1991.1; -;
 DR EMBL; BC017337; AAH17337.1; -;
 DR EMBL; BC011355; AAH11355.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 519 AA; 59570 MW; 088B8B6125CEEF36 CRC64;

Query Match 87.8%; Score 36; DB 4; Length 519;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
 |||||
 DB 70 YCHPVC 75

RESULT 7

Q8R2T8 PRELIMINARY; PRT; 520 AA.
 ID Q8R2T8
 AC Q8R2T8
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similar to general transcription factor IIIC, polypeptide 5
 DE (63KD).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC027247; AAH27247.1; -;
 SQ SEQUENCE 520 AA; 60471 MW; D8964D340B30F009 CRC64;

Query Match 87.8%; Score 36; DB 11; Length 520;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCHAVC 6
 |||||
 DB 70 YCHPVC 75

RESULT 8

Q96PV6 PRELIMINARY; PRT; 795 AA.
 ID Q96PV6
 AC Q96PV6
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE KIAA1932 protein (Fragment).
 GN KIAA1932
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=21456161; PubMed=11572484;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XXI.
 The complete sequences of 60 new cDNA clones from brain which code for
 large proteins.";
 RL DNA Res. 8:179-187(2001).
 DR EMBL; AB067519; BAB67825.1; -;
 DR InterPro; IPR005062; SAC3_GANP.
 DR Pfam; PF03399; SAC3_GANP; 1.
 FT NON_TER 1

SQ SEQUENCE 795 AA; 87903 MW; 2034E570B6535E66 CRC64;
Query Match 85.4%; Score 35; DB 4; Length 795;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 YCHAVC 6
DB 697 YCHAPC 702
RESULT 9
Q9D974 PRELIMINARY; PRT; 114 AA.
AC Q9D974
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE 1700127D068Rik protein.
GN 1700127D068Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kakuwaka T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Konteuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL, AK007298; BAB24941.1; -;
DR HSPF, F36368; IAO5.
DR MEROPS, S01.107; -;
DR MGD, MGI:1924249; 1700127D068Rik.
DR InterPro, IPR001230; Prenyl site.
DR InterPro, IPR001254; Ser protease_Try.
DR Pfam, PF00089; trypsin; 1.
DR SMART, SM00020; tryp_SPC; 1.
DR PROSITE, PS00294; PRENYLATION; UNKNOWN_1.
DR PROSITE, PS0240; TRYPSIN_DOM; 1.
DR PROSITE, PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 114 AA; 13082 MW; CF6C05A967C47546 CRC64;
Query Match 82.9%; Score 34; DB 11; Length 114;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 35 CHAVC 39
RESULT 10
Q91013

ID Q91013 PRELIMINARY; PRT; 170 AA.
AC Q91013
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TR110.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TB40E; AND 36YUEK;
RA Spaderna S., Blessing H., Bogner E., Britt W., Mach M.;
RT "Identification of glycoprotein gpTRL10 as a structural component of
RL human cytomegalovirus.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF432092; AAL27474.1; -;
DR EMBL; AF432086; AAL27468.1; -;
SQ SEQUENCE 170 AA; 18940 MW; B496EE5601E33739 CRC64;
Query Match 82.9%; Score 34; DB 12; Length 170;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 59 CHAVC 63
RESULT 11
Q8SBAL PRELIMINARY; PRT; 235 AA.
ID Q8SBAL
AC Q8SBAL
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 25.9 kDa protein.
GN OSUNBA0042H09.16.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Hsiao J., Ziemann V., Blunt S., Pui G.,
RA VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSUNBA0042H09 genomic sequence.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079874; AAL79784.1; -;
KW Hypothetical protein.
SQ SEQUENCE 235 AA; 25948 MW; 11D13F750B1046F7 CRC64;
Query Match 82.9%; Score 34; DB 10; Length 235;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 113 CHAVC 117
RESULT 12
Q9FPEI PRELIMINARY; PRT; 238 AA.
ID Q9FPEI
AC Q9FPEI
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 26.6 kDa protein (Fragment).
GN T14P1.19/AT2G45010.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayaishizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.
RT "Full length cDNA of gene T14P1.19/AT2G45010 (GI:4895249).";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327424; AAG42014.1;
KW Hypothetical protein.
FT NON_TER 238
SQ SEQUENCE 238 AA; 26581 MW; B272FAA3E46A7FB7 CRC64;
Query Match 82.9%; Score 34; DB 10; Length 238;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 106 CHAVC 110
|||||

RESULT 13
Q9SHD7
ID Q9SHD7 PRELIMINARY; PRT; 244 AA.
AC Q9SHD7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE At2G45010 protein.
DN "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana."
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buel C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Unayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana."
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007659; AAD32834.1;
SQ SEQUENCE 244 AA; 27264 MW; F03B26523376272 CRC64;
Query Match 82.9%; Score 34; DB 10; Length 244;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
DB 106 CHAVC 110
|||||

RESULT 14
Q9VH34
ID Q9VH34 PRELIMINARY; PRT; 380 AA.
AC Q9VH34
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
GN CG12812 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nufskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
DR EMBL; AE003685; AAP54486.1;
DR FlyBase; FBgn0037781; CG12812.
SQ SEQUENCE 380 AA; 43384 MW; 194DCE07A99B00C7 CRC64;
Query Match 82.9%; Score 34; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6
DB 318 CHAVC 322
|||||

RESULT 15

Q8T913 ID Q8T913 PRELIMINARY; PRT; 381 AA.
AC Q8T913;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE A707283p.
GN CG12812.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnikier S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY075174; AAL68044.1; -.
SQ SEQUENCE 381 AA; 43725 MW; 012595FB8830BC9C CRC64;
Query Match 82.9%; Score 34; DB 5; Length 381;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 338 CHAVC 342

RESULT 16

Q9FRE7 ID Q9FRE7 PRELIMINARY; PRT; 424 AA.
AC Q9FRE7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 46.7 kDa protein.
GN OSJNBAA0013M12.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatodeae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Ziemann V., Pai G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
RA Bowman C.L., Craven B., Uterback T.R., Khalak H., Feldblum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RL "Oryza sativa chromosome 3 BAC OSJNBAA0013M12 genomic sequence."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC082644; AAG46124.1; -.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
KW Hypothetical protein.
SQ SEQUENCE 424 AA; 46722 MW; 19B44CE09CF41F02 CRC64;
Query Match 82.9%; Score 34; DB 10; Length 424;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CHAVC 6
DB 39 CHAVC 43

RESULT 17

Q9MZUS ID Q9MZUS PRELIMINARY; PRT; 537 AA.
AC Q9MZUS;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Inter-cellular adhesion molecule-1 precursor.
GN ICAM-1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=DOMESTIC BREED; TISSUE=AOHTA;
RA Stocker C., Sugars K., Yarwood H., Delikouras A., Dorling A.,
RA Lechler R., Landis C., Morley B., Haskard D.;
RT "Cloning of porcine ICAM-1 and characterization of ITS induction on
RT endothelial cells by cytokines."
RL J. Immunol. 0:0-0(1999).
DR EMBL; AF156712; AAF80287.1; -.
DR HSP; P05362; 1IC1.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 3.
DR SMART; SM00409; IG; 1.
KW SIGNAL.
FT SIGNAL. 1 30 POTENTIAL.
FT CHAIN 31 537 INTERCELLULAR ADHESION MOLECULE-1.
SQ SEQUENCE 537 AA; 58430 MW; 47DA0F0F3C75CD54 CRC64;
Query Match 82.9%; Score 34; DB 6; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

DB 96 CHAVC 100

RESULT 18

Q96ZZO ID Q96ZZO PRELIMINARY; PRT; 646 AA.
AC Q96ZZO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein ST1701.
GN ST1701.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sakine M., Baba S.-I., Ankaei A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RL "Complete genome sequence of an aerobic thermophilic
RL Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000987; BAB66783.1; -.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR Pfam; PF00037; fer4; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; UNKNOWN 1.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 646 AA; 73755 MW; EC6C3547A9F8BE0F CRC64;

Query Match 82.9%; Score 34; DB 17; Length 646;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6
:||||:
DB 457 YCNAVC 462

RESULT 19

Q8T3A6 O8T3A6 PRELIMINARY; PRT; 1045 AA.
AC O8T3A6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Y47H9C.4c protein.
GN Y47H9C.4C.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AL032657; CAD27615.1; -;
SQ SEQUENCE 1045 AA; 111723 MW; 0A6D57A3A80BCDCA CRC64;

Query Match 82.9%; Score 34; DB 5; Length 1045;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6
:||||:
DB 847 FCHALC 852

RESULT 20

Q8T3A7 O8T3A7 PRELIMINARY; PRT; 1070 AA.
AC O8T3A7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Y47H9C.4b protein.
GN Y47H9C.4B.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AL032657; CAD27614.1; -;

SQ SEQUENCE 1070 AA; 114180 MW; 75254D0DD5643AE5 CRC64;

Query Match 82.9%; Score 34; DB 5; Length 1070;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6
:||||:
DB 847 FCHALC 852

RESULT 21

Q9XWD6 Q9XWD6 PRELIMINARY; PRT; 1111 AA.
AC Q9XWD6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Y47H9C.4 protein (CED-1).
GN Y47H9C.4 OR CED-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R.; Ainscough R.; Anderson K.; Baynes C.; Berks M.;
Bonfield J.; Burton J.; Connell M.; Copsey T.; Cooper J.; Coulson A.;
Craxton M.; Dear S.; Du Z.; Durbin R.; Favello A.; Fulton L.;
Gardner A.; Green P.; Hawkins T.; Hillier L.; Jier M.; Johnston L.;
Jones M.; Kershaw J.; Kirsten J.; Laister N.; Latreille P.;
Lightning J.; Lloyd C.; McMurray A.; Mortimore B.; O'Callaghan M.;
Parsons J.; Percy C.; Rifken L.; Roopra A.; Saunders D.; Showkneen R.;
Smailson N.; Smith A.; Sonhammer E.; Staden R.; Sulston J.;
Thierry-Mieg J.; Thomas K.; Vaudin M.; Vaughan K.; Waterston R.;
Watson A.; Weinstock L.; Wilkinson-Sproat J.; Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans";
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21097720; PubMed=11163239;
RA Zhou Z.; Hartwig E.; Horvitz H.R.;
RT "CED-1 is a Transmembrane Receptor that Mediates Cell Corpse
Engulfment in C. elegans";
RL Cell 104:43-56(2001).
DR EMBL; AL032657; CAA21739.1; -;
DR EMBL; AF332568; AAG60061.1; -;
DR HSSP; P05106; IJV2.1;
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 9.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00180; EGF Lam; 6.
DR SMART; SM00001; EGF-like; 5.
DR SMART; SM00261; FU; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 15.
DR PROSITE; PS01186; EGF_2; 11.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 1111 AA; 118803 MW; A39F374C008F9874 CRC64;

Query Match 82.9%; Score 34; DB 5; Length 1111;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCHAVC 6
:||||:
DB 847 FCHALC 852

Db 847 FCHALC 852

Query Match 82.9%; Score 34; DB 12; Length 3115;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 24

Q91FX0 PRELIMINARY; PRT; 3115 AA.

ID Q91FX0

AC Q91FX0

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE P349 protein.

OS Citrus tristeza virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;

OC Closterovirus.

OX NCBI_TaxID=12162;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=T30;

EX MEDLINE=20347348; PubMed=10888625;

RA Albiach-Marti M.R., Mawassi M., Gowda S., Satynaravana T., Hilf M.E.,
Shanker S., Almira E.C., Vives M.C., Lopez C., Guerri J., Flores R.,
Moreno P., Garnsey S.M., Dawson W.O.;
"Sequences of Citrus tristeza virus separated in time and space are
essentially identical.";
J. Virol. 74:6856-6865(2000).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=T30;

RA Albiach-Marti M.R., Dawson W.O.;
Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF260651; AAF70348.1; -

DR MEROPS; C42.001; -

DR InterPro; IPR000515; BPD transp.

DR InterPro; IPR000606; Viral helicase1.

DR InterPro; IPR002588; V.methyltransf.

DR Pfam; PF01443; Viral helicase1; 1.

DR Pfam; PF01660; Vmethyltransf; 1.

DR PROSITE; PS00402; BPD TRANSP INN MEMBER; UNKNOWN 1.

SQ SEQUENCE 3115 AA; 346872 MW; 6897F494ED25AD63 CRC64;

Query Match 82.9%; Score 34; DB 12; Length 3115;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

Db 600 CHAVC 604

RESULT 25

Q9WID7 PRELIMINARY; PRT; 3115 AA.

ID Q9WID7

AC Q9WID7

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE P349 protein.

OS Citrus tristeza virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;

OC Closterovirus.

OX NCBI_TaxID=12162;

RN [1]

RP SEQUENCE FROM N.A.

EX MEDLINE=99190444; PubMed=10092023;

RA Vives M.C., Rubio L., Lopez C., Navas-Castillo J., Albiach-Marti M.R.,
Dawson W.O., Guerri J., Flores R., Moreno P.;

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=SY568;

RC Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF001623; AAB57702.1; -

DR MEROPS; C42.001; -

DR InterPro; IPR000606; Viral helicase1.

DR InterPro; IPR002588; V.methyltransf.

DR Pfam; PF01443; Viral helicase1; 1.

DR Pfam; PF01660; Vmethyltransf; 1.

DR Polyprotein.

SQ SEQUENCE 3115 AA; 347070 MW; B406EAOFB38E73F9 CRC64;

Db 847 FCHALC 852

Query Match 82.9%; Score 34; DB 12; Length 3107;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

Db 594 CHAVC 598

RESULT 23

O10467 PRELIMINARY; PRT; 3115 AA.

ID O10467

AC O10467

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE 347-kDa polyprotein.

OS Citrus tristeza virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;

OC Closterovirus.

OX NCBI_TaxID=12162;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SY568;

EX MEDLINE=20007030; PubMed=10541017;

RA Yang Z.N., Mathews D.M., Dodds J.A., Mirkov T.E.;
"Molecular characterization of an isolate of citrus tristeza virus
that causes severe symptoms in sweet orange.";
Virus Genes 19:131-142(1999).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=SY568;

RC Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF001623; AAB57702.1; -

DR MEROPS; C42.001; -

DR InterPro; IPR000606; Viral helicase1.

DR InterPro; IPR002588; V.methyltransf.

DR Pfam; PF01443; Viral helicase1; 1.

DR Pfam; PF01660; Vmethyltransf; 1.

DR Polyprotein.

SQ SEQUENCE 3115 AA; 347070 MW; B406EAOFB38E73F9 CRC64;

RT "The complete genome sequence of the major component of a mild citrus
 RT tristeza virus isolate."
 RL J. Gen. Virol. 80:811-816(1999).
 RN [2]

RP SEQUENCE FROM N.A.

RA Moreno P.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; Y18420; CAA77161.1; -

DR MEROPS; C42.001; -

DR InterPro; IPR000515; BPD_transp.

DR InterPro; IPR000606; Viral_helicase1.

DR InterPro; IPR002588; V_methyltransf.

DR Pfam; PF01443; Viral_helicase1; 1.

DR Pfam; PF01660; Vmethyltransf; 1.

DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN 1.

SQ SEQUENCE 3115 AA; 346980 MW; 3EEA612F605D5045 CRC64;

Query Match

82.9%; Score 34; DB 12; Length 3115;

Best Local Similarity 100.0%; Pred. No. 4.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CHAVC 6

||||

Db 600 CHAVC 604

Search completed: June 20, 2003, 20:56:53

Job time : 36.6364 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 31.9318 Seconds
(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-12

Perfect score: 27

Sequence: 1 KHAVD 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	5	19	AAW23985
2	27	100.0	5	20	AA17120
3	27	100.0	5	20	AAW91034
4	27	100.0	5	21	AA197347
5	27	100.0	5	21	AAV73803
6	27	100.0	5	22	AAW65372
7	27	100.0	5	23	AAW47479
8	27	100.0	109	22	AAW05777
9	27	100.0	159	22	AAW91507
10	27	100.0	159	22	AAW79915

ALIGNMENTS

RESULT 1
AAW23985
ID

AAW23985 standard; peptide; 5 AA.

11	27	100.0	159	22	AAW79916
12	27	100.0	181	21	AAW31690
13	27	100.0	211	21	AAW31689
14	27	100.0	214	21	AAW42106
15	27	100.0	214	21	AAW43503
16	27	100.0	237	21	AAW42105
17	27	100.0	237	21	AAW43502
18	27	100.0	265	21	AAW42104
19	27	100.0	265	21	AAW43501
20	27	100.0	274	22	AAW63327
21	27	100.0	274	20	AAW23782
22	27	100.0	478	20	AAW73971
23	27	100.0	478	21	AAW23259
24	27	100.0	478	21	AAW12885
25	27	100.0	478	22	AAW78669
26	27	100.0	605	21	AAW12884
27	27	100.0	605	22	AAW78668
28	27	100.0	659	23	AAW23571
29	27	100.0	727	22	AAW60064
30	27	100.0	730	22	AAW61435
31	27	100.0	1655	22	AAW61965
32	27	100.0	1856	20	AAW21801
33	27	100.0	1856	21	AAW83269
34	27	100.0	2868	22	AAW79895
35	26	96.3	75	18	AAW55618
36	26	96.3	92	18	AAW55343
37	26	96.3	255	23	ABP30480
38	26	96.3	272	22	AAW82359
39	26	96.3	283	23	ABP27879
40	26	96.3	295	23	ABP39397
41	26	96.3	296	22	ABW60556
42	26	96.3	317	22	ABW71981
43	26	96.3	425	22	ABW58748
44	26	96.3	429	22	AAU37298
45	26	96.3	429	22	AAU37516
46	26	96.3	492	22	AAU34312
47	26	96.3	492	22	AAW00839
48	26	96.3	494	20	AAW85670
49	26	96.3	664	23	ABW47519
50	26	96.3	1506	23	AAW47585
51	26	96.3	4150	21	AAW92707
52	26	96.3	4150	21	AAW16307
53	24	88.9	38	22	AAW11633
54	24	88.9	42	18	AAW16035
55	24	88.9	56	20	AAW60375
56	24	88.9	57	21	AAW21434
57	24	88.9	76	22	AAW48134
58	24	88.9	77	23	ABP09660
59	24	88.9	83	19	AAW77330
60	24	88.9	115	23	ABP31852
61	24	88.9	130	21	AAW64889
62	24	88.9	152	21	AAW62888
63	24	88.9	159	20	AAW99659
64	24	88.9	168	22	AAW13564
65	24	88.9	171	21	AAW43698
66	24	88.9	177	23	ABW97300
67	24	88.9	191	20	AAW66159
68	24	88.9	191	23	ABW79058
69	24	88.9	191	23	AAE19601
70	24	88.9	191	23	AAE18886
71	24	88.9	191	23	ABW08709
72	24	88.9	194	20	AAW60053
73	24	88.9	199	21	AAW73435
74	24	88.9	199	22	ABW58473
75	24	88.9	199	22	AAW994259

RESULT 1
AAW23985
ID

AAW23985 standard; peptide; 5 AA.

XX AC AAW23985;
 XX 23-JUL-1998 (first entry)
 XX Cadherin-mediated cell adhesion modulating cyclic peptide 7.
 XX
 XX Cadherin-mediated; cell adhesion; drug delivery; treatment; cancer;
 XX angiogenesis; skin grafting; neurite growth; veterinary medicine;
 KW wound healing; tumour; metastasis; carcinoma; leukaemia; modulation;
 KW organ transplant; neurological disease; cyclic.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "optional N-terminal acetylation;
 FT amide bond formed between the residue
 FT side chains of Lys at this position
 FT and Asp at position 5"
 FT
 FT Modified-site 5 /note= "optional C-terminal amidation;
 FT amide bond formed between the residue
 FT side chains of Asp at this position
 FT and Lys at position 1"
 XX
 XX WO9802452-A2.
 XX
 XX 22-JAN-1998.
 XX
 XX 11-JUL-1997; 97WO-CA00489.
 XX
 XX 12-JUL-1996; 96US-0021612.
 XX (UYMC-) UNIV MCGILL.
 XX
 XX Blaschuk OW, Gour BJ;
 XX WPI; 1998-110522/10.
 XX
 XX Cyclic peptide(s) that modulate cadherin-mediated cell adhesion -
 XX used to improve drug delivery through skin, to the brain etc., for
 XX treatment of cancer, angiogenesis etc., and to improve grafting of
 XX foreign tissue or neurite growth
 XX
 XX Claim 17; Page 98; 133pp; English.
 XX
 XX This is a cyclic peptide capable of modulating cadherin-mediated cell
 XX adhesion. Cadherin-expressing cells can be detected from their reaction
 XX with the antibody that binds the peptides of the invention. These
 XX peptides are useful in human or veterinary medicine to modulate adhesion
 XX mediated by E-, N-, P- or R-cadherins (or other cadherins that include
 XX the amino acids HAV recognition sequence), specifically where these are
 XX expressed by epithelial, endothelial, neural or tumour cells or
 XX lymphocytes. The peptides which inhibit cell adhesion are used to improve
 XX delivery of drugs through the skin (such that the peptides enters the
 XX blood stream) and to tumours (particularly ovarian or bladder tumours or
 XX melanoma) and to the brain. They are also used to treat cancer
 XX (carcinoma, leukaemia or melanoma), inhibit metastasis and also inhibit
 XX angiogenesis. The peptides that stimulate adhesion are used to improve
 XX wound healing to promote adherence of foreign tissues (skin grafts or
 XX organ transplants), and to improve adherence to tissue culture surfaces
 XX and bioeffectors. Other uses of the peptides are to induce apoptosis in
 XX cadherin-expressing cells, increase or decrease neurite outgrowth, to
 XX treat spinal cord injuries and de-myelinating neurological diseases
 XX (specifically multiple sclerosis). They can be used for modulating the
 XX immune system (e.g. in cases of diabetes or rheumatoid arthritis), for
 XX preventing pregnancy and to increase vaso-permeability. The antibodies
 XX which bind to these peptides are also used to modulate cell adhesion and
 XX when coupled to a drug, to target the drug to cadherin-expressing cells.
 XX Diagnostic agents may also be delivered using the peptides in a skin
 XX patch, particularly where the agent generates a colour in contact with,
 XX e.g. cocaine, human immunodeficiency viral proteins, glucose or

CC prostate-specific antigen, particularly for home-testing kits. The
 CC peptides make it possible to deliver a wide range of drugs through the
 CC skin, avoiding parenteral administration, by-passing the gastro-
 CC intestinal system and improving patient compliance.
 XX
 XX Sequence 5 AA;
 XX
 XX Query Match 100.0%; Score 27; DB 19; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 KHAVD 5
 XX |||||
 XX Db 1 KHAVD 5
 XX
 XX RESULT 2
 XX AAY17120
 XX ID AAY17120 standard; peptide; 5 AA.
 XX AC AAY17120;
 XX XX 20-SEP-1999 (first entry)
 XX XX Cadherin-mediated adhesion modulating cyclic peptide.
 XX XX Cell adhesion modulation; CAM; synaptic stability; cadherin;
 XX KW cadherin-mediated adhesion; drug delivery; cell adhesion; tumour;
 XX KW wound healing; neurite outgrowth; cyclic; circular.
 XX OS Synthetic.
 XX XX WO9933875-A1.
 XX XX 08-JUL-1999.
 XX XX 23-DEC-1998; 98WO-CA01207.
 XX XX 23-DEC-1997; 97US-0996679.
 XX XX (UYMC-) UNIV MCGILL.
 XX XX Blaschuk OW, Gour BJ;
 XX XX WPI; 1999-430231/36.
 XX XX Cyclic peptide cell adhesion modulating agents, useful for
 XX PT modulating synaptic stability
 XX PS Claim 17; Page 65; 144pp; English.
 XX XX The invention provides cyclic peptide cell adhesion modulating (CAM)
 XX agents that comprises a His-Ala-Val recognition sequence. Also provided
 XX is a method for inhibiting synaptic stability in a mammal that comprises
 XX administering to a mammal a therapeutically effective amount of a CAM
 XX agent that inhibits cadherin-mediated adhesion, where the agent comprises
 XX a cyclic peptide having a peptide ring, and where the sequence His-Ala-
 XX Val is present within the peptide ring. The cyclic peptides are cell
 XX adhesion modulating agents that inhibit cadherin-mediated adhesion. They
 XX can be used in a method for inhibiting synaptic stability in mammals. The
 XX agents can be used to treat diseases or other conditions characterized by
 XX undesirable cell adhesion or to facilitate drug delivery to a specific
 XX tissue or tumour. Alternatively the agents may be used to enhance cell
 XX adhesion (e.g. to supplement or replace stitches or to facilitate wound
 XX healing) or to enhance or direct neurite outgrowth.
 XX
 XX Sequence 5 AA;
 XX
 XX Query Match 100.0%; Score 27; DB 20; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 KHAVD 5

Db 1 KHAVD 5

RESULT 3

AAW91034
ID AAW91034 standard; peptide; 5 AA.

XX AC AAW91034;

XX DT 24-MAR-1999 (first entry)

XX DE Peptide inhibiting interaction of alpha-catenin and beta-catenin.

XX KW Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;
KW cadherin-mediated function; demyelinating neurological disease;
KW multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;
KW central nervous system; apoptosis induction; cadherin-expression cell;
KW pregnancy prevention; vasopermeability; synaptic stability; diabetes;
KW rheumatoid arthritis; allergic response; learning; memory;
KW antibody-mediated graft rejection; cyclic.

XX OS Synthetic.

XX PN WO9845319-A2.

XX PD 15-OCT-1998.

XX PF 14-APR-1998; 98WO-CA00322.

XX PR 10-APR-1997; 97US-0043361.

XX PA (UYMC-) UNIV MCGILL.

XX PI Blaschuk OW, Gour BJ;

XX DR WPI; 1999-024009/02.

XX PT New catenin modulating agents - comprising peptides having a
PT sequence HAV or analogues or antibodies, used for modulating
PT cadherin-mediated functions

XX PS Claim 12; Page 78; 106pp; English.

XX CC The present sequence represents a peptide which is capable of inhibiting
CC an interaction between alpha-catenin and beta-catenin. The peptide is
CC used in modulating agents that are used for modulating cadherin-mediated
CC functions. They can be used for disrupting interaction between
CC alpha-catenin and beta-catenin in a cell, inhibiting cell adhesion, e.g.
CC between epithelial cells, endothelial cells, neural cells, tumour cells
CC and lymphocytes, for treating a demyelinating neurological disease, e.g.
CC multiple sclerosis, for reducing unwanted cellular adhesion in a mammal,
CC for enhancing the delivery of a drug through the skin of a mammal, for
CC enhancing the delivery of a drug to a tumour in a mammal, for treating
CC cancer in a mammal, for inhibiting angiogenesis in a mammal, for
CC enhancing drug delivery to the central nervous system of a mammal, for
CC inducing apoptosis in a cadherin-expression cell, for modulating the
CC immune system of a mammal, for preventing pregnancy in a mammal, for
CC increasing vasopermeability in a mammal, or for inhibiting synaptic
CC stability in a mammal. In particular they can be used for treating
CC diabetes, rheumatoid arthritis, allergic responses, antibody-mediated
CC graft rejection or for stimulating learning and memory.

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 20; Length 5;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5

Db 1 KHAVD 5

RESULT 5

AAW73803
ID AAW73803 standard; peptide; 5 AA.

XX XX AAW73803;

RESULT 4

AAB27347
ID AAB27347 standard; Peptide; 5 AA.

XX AC AAB27347;

XX DT 15-FEB-2001 (first entry)

XX DE Beta-catenin derived cyclic peptide SEQ ID NO: 35.

XX KW Beta-catenin; cadherin-mediated intercellular adhesion;
KW cell differentiation; modulating agent; hair loss; skin exfoliation;
KW internalisation moiety; flanking sequence; transcription; hearing loss;
KW cyclic.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1..5

XX FT /note= "the molecule is cyclised"

XX PN WO200053632-A1.

XX PD 14-SEP-2000.

XX PF 07-MAR-2000; 2000WO-CA00222.

XX PR 09-MAR-1999; 99US-0265107.

XX PA (UYMC-) UNIV MCGILL.

XX PI Blaschuk OW, Gour BJ;

XX DR WPI; 2000-594308/56.

XX PT Stimulating beta-catenin mediated gene expression, cellular
PT differentiation and hair growth, involves contacting cells with
PT modulating agent capable of inhibiting interaction between alpha and
PT beta catenin

XX PS Disclosure; Page 7; 77pp; English.

XX CC The present invention is concerned with methods of modulating the amount
CC of free beta-catenin in the cell, and methods of stimulating the
CC expression of genes involved in cellular differentiation, the
CC transcription of which is under the control of beta-catenin. The
CC peptides given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351
CC can be used as modulating agents which interrupt the interaction between
CC alpha and beta catenin, causing increased levels of the Wntter and
CC stimulating the activation of beta-catenin mediated transcription. This
CC can be used to stimulate cell differentiation, which can then be used to
CC promote hair growth and skin exfoliation. This latter is particularly
CC useful in the improvement of photodamaged skin and to minimise wrinkles.
CC The modulating peptide can also be used to reduce hearing loss resulting
CC from inner ear disorders such as hyperacusis and tinnitus.

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 21; Length 5;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5

Db 1 KHAVD 5

XX DT 07-MAR-2000 (first entry)
 XX DE N-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:309.
 XX KW Diagnosis; cancer; OB-cadherin; N-cadherin; E-cadherin; CAR; cyclic;
 KW cell adhesion recognition sequence; leukaemia; metastasis;
 KW prostate cancer; breast cancer; ovarian cancer; carcinoma.
 XX OS Synthetic.
 XX OS Homo sapiens.
 XX PN WO9957565-A2.
 XX PD 11-NOV-1999.
 XX PF 05-MAY-1999; 99WO-CA00362.
 XX PR 05-MAY-1998; 98US-0073040.
 XX PR 06-NOV-1998; 98US-0187859.
 XX PR 20-JAN-1999; 99US-0234395.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Blaschuk OW, Gour BJ, Byers S;
 XX DR WPI; 2000-062165/05.
 XX PT Detecting expression of OB-cadherin or N-cadherin used for diagnosing
 XX PT and evaluating cancer -
 XX PS Disclosure; Page 22; 56pp; English.
 XX CC The present invention describes methods which have been developed for
 CC detecting and evaluating cancer that are based on the finding that
 CC OB-cadherin and N-cadherin are expressed by metastatic carcinoma cells,
 CC but not by highly differentiated, poorly invasive carcinomas. A method
 CC for determining the presence or absence of a cancer in a patient
 CC comprises: (1) contacting a biological sample from the patient with a
 CC binding agent that specifically binds to OB- or N-cadherin OR with an
 CC oligonucleotide that hybridizes to a polynucleotide that encodes OB- or
 CC N-cadherin; and (2) detecting in the sample an amount of polypeptide
 CC that binds to the binding agent OR the amount of polynucleotide that
 CC hybridizes to the oligonucleotide, relative to a predetermined cut-off
 CC value, and determining the presence or absence of cancer in the patient
 CC from this. The methods from the present invention can be used to
 CC determine the metastatic potential of a cancer. The methods may be used
 CC to detect a metastatic cancer in a patient, to monitor progression of a
 CC cancer, or to evaluate the metastatic potential of a cancer. Cancers
 CC which may be evaluated using the methods include leukemia, prostate
 CC cancer, breast cancer and ovarian cancer. AAY73501 to AAY73813 represent
 CC peptide sequences used in the exemplification of the present invention.
 CC Specifically, AAY73503 to AAY73789 represent OB-cadherin cell adhesion
 CC recognition (CAR) peptides and AAY73790 to AAY73808 represent N-cadherin
 CC CAR peptides. AAY47906 to AAY47915 represent PCR primers used in examples
 CC from the present invention.
 XX SQ Sequence 5 AA;
 Query Match 100.0%; Score 27; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVD 5
 Db 1 KHAVD 5
 RESULT 6
 AAG65372
 ID AAG65372 standard; peptide; 5 AA.
 XX AC AAG65372;

XX DT 30-NOV-2001 (first entry)
 XX DE Cyclic peptide with cadherin CAR sequence.
 XX KW Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 KW cytoskeletal; vulnary; immunomodulator; vasotropic; neuroprotective;
 KW cerebroprotective; muscular; cyclic.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetylation"
 FT Modified-site 5 /note= "C-terminal amide"
 FT WO200153331-A2.
 XX PN 26-JUL-2001.
 XX PD 24-JAN-2001; 2001WO-US02508.
 XX PF 24-JAN-2000; 2000US-0491078.
 XX PR (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PA Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 XX PI Hu Z;
 XX DR WPI; 2001-549899/61.
 XX PT Cell adhesion modulating agent used for enhancing delivery of drug to
 XX PT tumor comprises imidazole compounds -
 XX PS Examples; Fig 3A-I; 436pp; English.
 XX CC The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cyclic peptide with
 CC cadherin cell adhesion recognition (CAR) sequence.
 XX SQ Sequence 5 AA;
 Query Match 100.0%; Score 27; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVD 5
 Db 1 KHAVD 5
 RESULT 7
 AAM47479
 ID AAM47479 standard; peptide; 5 AA.
 XX AC AAM47479;
 XX DT 12-FEB-2002 (first entry)
 XX DE Cyclic peptide endothelial cell adhesion modulator #2.
 XX AC

PR 03-AUG-2000; 2000JP-0280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
DR N-PSDB; AAH66726.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
XX Claim 17; SEQ ID NO: 5261; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 159 AA;

Query Match 100.0%; Score 27; DB 22; Length 159;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db |||||
36 KHAVD 40

RESULT 10
AAB79915
ID AAB79915 standard; Protein; 159 AA.
XX
AC AAB79915;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:564.
XX
XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
FN WO200100843-A2.
XX
XX 04-JAN-2001.
PD
XX 23-JUN-2000; 2000WO-IB00923.
PF
XX 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 02-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.

PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 08-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032186.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032230.
PR 09-JUL-1999; 99DE-1032322.
PR 14-JUL-1999; 99DE-1032326.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 31-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.

(SADI) BASF AG.

Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;

WPI; 2001-137957/14.
N-PSDB; AAF72034.

Nucleic acids from Corynebacterium glutamicum encoding metabolic
pathway proteins, useful for producing fine chemicals in
microorganisms, including organic acids, nonproteinogenic amino acids,
and purine and pyrimidine bases -

Claim 20; Page 963-964; 1737pp; English.

AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
pathway (MP) proteins given in AAB79634 to AAB80241. The C. glutamicum
MP nucleic acids are useful for the production of fine chemicals
in microorganisms, including organic acids, nonproteinogenic amino
acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
compounds, vitamins, cofactors, polyketides and enzymes.

```
XX SQ Sequence 159 AA;
Query Match 100.0%; Score 27; DB 22; Length 159;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 36 KHAVD 40

RESULT 11
AAB79916
ID AAB79916 standard; Protein; 159 AA.
AC AAB79916;
XX
XX 30-APR-2001 (first entry)
XX
XX Corynebacterium glutamicum MP protein sequence SEQ ID NO:566.
XX
XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
XX WO200100843-A2.
XX
XX 04-JAN-2001.
XX
XX 23-JUN-2000; 2000WO-IB00923.
XX
XX 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 02-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032186.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032926.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.

XX SQ Sequence 159 AA;
Query Match 100.0%; Score 27; DB 22; Length 159;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 36 KHAVD 40

RESULT 12
AAG31690
ID AAG31690 standard; Protein; 181 AA.
AC AAG31690;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38099.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
```

PD	06-SEP-2000.		PR	06-JUL-1999;	99US-0142390.
XX			PR	08-JUL-1999;	99US-0142803.
PF	25-FEB-2000;	2000EP-0301439.	PR	09-JUL-1999;	99US-0142920.
XX			PR	12-JUL-1999;	99US-0142977.
PR	25-FEB-1999;	99US-0121825.	PR	13-JUL-1999;	99US-0143542.
PR	05-MAR-1999;	99US-0123180.	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	99US-0123548.	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	99US-0125788.	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	99US-0126264.	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	99US-0126785.	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	99US-0127462.	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	99US-0128234.	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	99US-0128714.	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	99US-0129845.	PR	19-JUL-1999;	99US-0144334.
PR	13-APR-1999;	99US-0130077.	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	30-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	05-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	06-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134219.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134321.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	21-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139452.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	21-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139753.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140921.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.

```

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160983.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.08; Score 27; DB 21; Length 181;
Best Local Similarity 100.08; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 139 KHAVD 143

RESULT 13
ID AAG31689 standard; Protein; 211 AA:
XX
AC AAG31689;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38098.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

```

```
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148694.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 169 KHAVD 173

RESULT 14
AAG42106
ID AAG42106 standard; Protein; 214 AA.
XX AC AAG42106;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 52470.14
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
```


PR	19-JUL-1999;	99US-01443339;
PR	20-JUL-1999;	99US-01443352;
PR	20-JUL-1999;	99US-01446332;
PR	20-JUL-1999;	99US-01448484;
PR	21-JUL-1999;	99US-01448414;
PR	21-JUL-1999;	99US-01450866;
PR	21-JUL-1999;	99US-01450888;
PR	22-JUL-1999;	99US-01450885;
PR	22-JUL-1999;	99US-01450887;
PR	22-JUL-1999;	99US-01450889;
PR	22-JUL-1999;	99US-01451192;
PR	23-JUL-1999;	99US-01451454;
PR	23-JUL-1999;	99US-01452118;
PR	23-JUL-1999;	99US-01452224;
PR	26-JUL-1999;	99US-01452766;
PR	27-JUL-1999;	99US-01459113;
PR	27-JUL-1999;	99US-01459118;
PR	27-JUL-1999;	99US-01459139;
PR	28-JUL-1999;	99US-01459519;
PR	02-AUG-1999;	99US-01463386;
PR	02-AUG-1999;	99US-01463388;
PR	02-AUG-1999;	99US-01463389;
PR	03-AUG-1999;	99US-01470338;
PR	04-AUG-1999;	99US-01472004;
PR	04-AUG-1999;	99US-01473002;
PR	05-AUG-1999;	99US-01471192;
PR	05-AUG-1999;	99US-01472660;
PR	06-AUG-1999;	99US-01473003;
PR	06-AUG-1999;	99US-01474166;
PR	09-AUG-1999;	99US-01474993;
PR	09-AUG-1999;	99US-01479935;
PR	10-AUG-1999;	99US-01481711;
PR	11-AUG-1999;	99US-01483319;
PR	12-AUG-1999;	99US-01483431;
PR	13-AUG-1999;	99US-01485655;
PR	13-AUG-1999;	99US-01486884;
PR	16-AUG-1999;	99US-01493688;
PR	17-AUG-1999;	99US-01491715;
PR	18-AUG-1999;	99US-01494266;
PR	20-AUG-1999;	99US-01497222;
PR	20-AUG-1999;	99US-01497231;
PR	20-AUG-1999;	99US-01499929;
PR	23-AUG-1999;	99US-01499902;
PR	23-AUG-1999;	99US-01499930;
PR	25-AUG-1999;	99US-01505666;
PR	26-AUG-1999;	99US-01508084;
PR	27-AUG-1999;	99US-01510065;
PR	27-AUG-1999;	99US-01510066;
PR	27-AUG-1999;	99US-01510080;
PR	30-AUG-1999;	99US-01513033;
PR	31-AUG-1999;	99US-01514338;
PR	01-SEP-1999;	99US-01519303;
PR	07-SEP-1999;	99US-01520363;
PR	10-SEP-1999;	99US-01530700;
PR	13-SEP-1999;	99US-01537558;
PR	15-SEP-1999;	99US-01540118;
PR	16-SEP-1999;	99US-01540339;
PR	20-SEP-1999;	99US-01547719;
PR	22-SEP-1999;	99US-01551339;
PR	23-SEP-1999;	99US-01554866;
PR	24-SEP-1999;	99US-01556599;
PR	28-SEP-1999;	99US-01564596;
PR	28-SEP-1999;	99US-01571117;
PR	04-OCT-1999;	99US-01571753;
PR	05-OCT-1999;	99US-01577865;
PR	06-OCT-1999;	99US-01580029;
PR	07-OCT-1999;	99US-01582022;
PR	12-OCT-1999;	99US-01583629;
PR	13-OCT-1999;	99US-01592393;
PR	13-OCT-1999;	99US-01592594;
PR	13-OCT-1999;	99US-01592595;
PR	14-OCT-1999;	99US-01593329;

Qy	Db	Query Match	Best Local Similarity	Score 27; DB 21; Length 214;	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
14-OCT-1999;	99US-0159330.				
14-OCT-1999;	99US-0159331.				
14-OCT-1999;	99US-0159637.				
14-OCT-1999;	99US-0159638.				
18-OCT-1999;	99US-0159584.				
21-OCT-1999;	99US-0160741.				
21-OCT-1999;	99US-0160767.				
21-OCT-1999;	99US-0160768.				
21-OCT-1999;	99US-0160770.				
21-OCT-1999;	99US-0160814.				
21-OCT-1999;	99US-0160815.				
22-OCT-1999;	99US-0160980.				
22-OCT-1999;	99US-0160981.				
22-OCT-1999;	99US-0160989.				
25-OCT-1999;	99US-0161404.				
25-OCT-1999;	99US-0161405.				
25-OCT-1999;	99US-0161406.				
26-OCT-1999;	99US-0161359.				
26-OCT-1999;	99US-0161360.				
26-OCT-1999;	99US-0161361.				
28-OCT-1999;	99US-0161920.				
28-OCT-1999;	99US-0161992.				
28-OCT-1999;	99US-0161993.				
29-OCT-1999;	99US-0162142.				
1 KHAVD 5					
12 KHAVD 16					
RESULT 15					
AAG43503					
ID AAG43503 standard; Protein; 214 AA.					
XX AC AAG43503;					
XX AC					
XX DT 18-OCT-2000 (first entry)					
XX XX					
XX XX Arabidopsis thaliana protein fragment SEQ ID NO: 54383.					
XX XX Protein identification; signal transduction pathway; metabolic pathway;					
KW KW hybridisation assay; genetic mapping; gene expression control; promoter;					
KW KW termination sequence.					
XX XX					
OS OS Arabidopsis thaliana.					
XX XX					
XX PN EPI033405-A2.					
XX XX					
PD PD 06-SEP-2000.					
XX XX					
XX XX 25-FEB-2000; 2000EP-0301439.					
XX XX					
PR PR 25-FEB-1999; 99US-0121825.					
PR PR 05-MAR-1999; 99US-0123180.					
PR PR 09-MAR-1999; 99US-0123548.					
PR PR 23-MAR-1999; 99US-0125788.					
PR PR 25-MAR-1999; 99US-0126264.					
PR PR 29-MAR-1999; 99US-0126785.					
PR PR 01-APR-1999; 99US-0127462.					
PR PR 06-APR-1999; 99US-0128234.					
PR PR 08-APR-1999; 99US-0128714.					
PR PR 16-APR-1999; 99US-0129845.					
PR PR 19-APR-1999; 99US-0130077.					
PR PR 21-APR-1999; 99US-0130449.					
PR PR 23-APR-1999; 99US-0130510.					
PR PR 23-APR-1999; 99US-0130891.					
PR PR 28-APR-1999; 99US-0131449.					
PR PR 30-APR-1999; 99US-0132047.					
PR PR 30-APR-1999; 99US-0132407.					
04-MAY-1999;	99US-0132484.				
PR PR 05-MAY-1999;	99US-0132485.				
PR PR 06-MAY-1999;	99US-0132486.				
PR PR 07-MAY-1999;	99US-0132487.				
PR PR 07-MAY-1999;	99US-0132863.				
PR PR 11-MAY-1999;	99US-0134256.				
PR PR 14-MAY-1999;	99US-0134218.				
PR PR 14-MAY-1999;	99US-0134219.				
PR PR 14-MAY-1999;	99US-0134221.				
PR PR 14-MAY-1999;	99US-0134370.				
PR PR 18-MAY-1999;	99US-0134768.				
PR PR 19-MAY-1999;	99US-0134941.				
PR PR 20-MAY-1999;	99US-0135124.				
PR PR 21-MAY-1999;	99US-0135353.				
PR PR 24-MAY-1999;	99US-0135629.				
PR PR 25-MAY-1999;	99US-0136021.				
PR PR 27-MAY-1999;	99US-0136392.				

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 214;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KHAVD 5
Db 12 KHAVD 16

RESULT 16

AAG42105
ID AAG42105 standard; Protein; 237 AA.

XX AC AAG42105;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 52469.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132407.

XX PR 05-MAY-1999; 99US-0132484.

XX PR 06-MAY-1999; 99US-0132485.

XX PR 07-MAY-1999; 99US-0132486.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 14-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

```
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 35 KHAVD 39

RESULT 17
AAG43502
ID AAG43502 standard; Protein; 237 AA.
XX AC AAG43502;
XX XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment. SEQ ID NO: 54382.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 200DEP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142330.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
```

```
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147192.
PR 07-AUG-1999; 99US-0147260.
PR 08-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 10-AUG-1999; 99US-0147493.
PR 11-AUG-1999; 99US-0147935.
PR 12-AUG-1999; 99US-0148171.
PR 13-AUG-1999; 99US-0148319.
PR 14-AUG-1999; 99US-0148341.
PR 15-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 19-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 21-AUG-1999; 99US-0149723.
PR 22-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 24-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 28-AUG-1999; 99US-0151066.
PR 29-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 02-SEP-1999; 99US-0152363.
PR 03-SEP-1999; 99US-0153070.
PR 04-SEP-1999; 99US-0153758.
PR 05-SEP-1999; 99US-0154018.
PR 06-SEP-1999; 99US-0154039.
PR 07-SEP-1999; 99US-0154779.
PR 08-SEP-1999; 99US-0155139.
PR 09-SEP-1999; 99US-0155486.
PR 10-SEP-1999; 99US-0155559.
PR 11-SEP-1999; 99US-0156458.
PR 12-SEP-1999; 99US-0156596.
PR 13-SEP-1999; 99US-0157117.
PR 14-SEP-1999; 99US-0157753.
PR 15-SEP-1999; 99US-0157865.
PR 16-SEP-1999; 99US-0158029.
PR 17-SEP-1999; 99US-0158232.
PR 18-SEP-1999; 99US-0158369.
PR 19-SEP-1999; 99US-0159293.
PR 20-SEP-1999; 99US-0159294.
PR 21-SEP-1999; 99US-0159295.
PR 22-SEP-1999; 99US-0159329.
PR 23-SEP-1999; 99US-0159330.
PR 24-SEP-1999; 99US-0159331.
PR 25-SEP-1999; 99US-0159337.
PR 26-SEP-1999; 99US-0159638.
PR 27-SEP-1999; 99US-0159638.
PR 28-SEP-1999; 99US-0160741.
PR 29-SEP-1999; 99US-0160767.
PR 30-SEP-1999; 99US-0160770.
PR 01-OCT-1999; 99US-0160814.
PR 02-OCT-1999; 99US-0160815.
PR 03-OCT-1999; 99US-0160980.
PR 04-OCT-1999; 99US-0160981.
PR 05-OCT-1999; 99US-0160989.
PR 06-OCT-1999; 99US-0161404.
PR 07-OCT-1999; 99US-0161405.
PR 08-OCT-1999; 99US-0161406.
PR 09-OCT-1999; 99US-0161359.
PR 10-OCT-1999; 99US-0161360.
PR 11-OCT-1999; 99US-0161361.
PR 12-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0161993.
PR 30-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 35 KHAVD 39

RESULT 18
AAG42104
ID AAG42104 standard; Protein; 265 AA.
XX AC AAG42104;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 52468.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
```

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 63 KHAVD 67

RESULT 19
AAG43501
ID AAG43501 standard; Protein; 265 AA.
XX AC AAG43501;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 54381.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.


```
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 63 KHAVD 67

RESULT 20

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 63 KHAVD 67

RESULT 20
```

```
AAG63327
ID AAG63327 standard; Protein; 274 AA.
XX
AC AAG63327;
XX
DT 01-OCT-2001 (first entry)
XX
DE Amino acid sequence of Apoptosis Related protein (ARTS).
XX
KW Apoptosis Related protein; ARTS; TGF-beta signalling pathway; septin;
XX gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155209-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02466.
XX
PR 29-JAN-2000; 2000US-0178866.
PR 29-DEC-2000; 2000US-0258725.
XX
(LARI/) LARISCH-BLOCH S.
PA (KIMS/) KIM S.
PA (LECH/) LECHLEIDER R J.
PA (ROBE/) ROBERTS A B.
PA (YIYY/) YI Y.
XX
Larisch-Bloch S, Kim S, Lechleider RJ, Roberts AB, Yi Y;
WPI; 2001-476191/51.
DR N-PSDB; AAH42965.
XX
New nucleic acids encoding an Apoptosis Related protein (ARTS) in the
TGF-beta signalling pathway for diagnosing apoptosis in a cell -
Claim 11; Page 6-7; 72pp; English.
XX
The present sequence represents an Apoptosis Related protein (ARTS) in
the TGF-beta signalling pathway. The ARTS protein is a member of the
septin family of proteins. ARTS is localised in the cytoplasm, in
particular the mitochondria. The ARTS protein contains a P-loop motif.
The C-terminal of ARTS appears to play a role in the cellular
localisation of ARTS. The presence of the ARTS protein in a cell is
indicative of apoptosis. ARTS polynucleotides can be used in gene
therapy protocols, and for the production of transgenic animals.
XX
SQ Sequence 274 AA;
Query Match 100.0%; Score 27; DB 22; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVD 5
DB 171 KHAVD 175
RESULT 21
AAY23782
ID AAY23782 standard; Protein; 478 AA.
XX
AC AAY23782;
XX
DT 13-SEP-1999 (first entry)
XX
DE Human cell division regulator (HCDR) 1.
XX
KW Human cell division regulator; HCDR; interphase; inflammation;
KW cell proliferation; apoptosis; neurodegenerative disease; aplastic anaemia; ischaemic injury;
KW neurodegenerative disease; viral infection; hepatitis B; hepatitis C.
KW liver damage; viral infection; hepatitis B; hepatitis C.
```

XX OS Homo sapiens.
 XX PN US928899-A.
 XX PD 27-JUL-1999.
 XX PF 01-OCT-1998; 98US-0165234.
 XX PR 15-OCT-1997; 97US-0951148.
 XX PR 01-OCT-1998; 98US-0165234.
 XX PA (INCY-) INCYTE PHARM INC.
 XX PI Bandman O, Corley NC, Hillman JL, Lal P, Shah P;
 XX DR WPI; 1999-429499/36.
 XX DR N-PSDB; AAX85940.
 XX PT Cell division regulators active in interphase
 XX PS Example 1; Fig 1A-E; 59pp; English.
 XX CC The present sequence represents human cell division regulator (HCDR) 1.
 CC HCDR proteins are active in interphase, and are used for the
 CC treatment or prevention of inflammation and disorders associated with
 CC cell proliferation and apoptosis. HCDR may be administered to a
 CC patient having a disorder associated with an increase in apoptosis,
 CC such a disorder may be e.g. neurodegenerative, a neurodegenerative
 CC disease, aplastic anaemia, an ischaemic injury, liver damage, or a
 CC viral infection such as hepatitis B or C.
 XX SQ Sequence 478 AA;
 Query Match 100.0%; Score 27; DB 20; Length 478;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 -Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVD 5
 DB 190 KHAVD 194
 |||||
 RESULT 22
 AAW73971
 ID AAW73971 standard; Protein; 478 AA.
 XX AC AAW73971;
 XX DT 04-MAY-1999 (first entry)
 XX DE Human HCDR-1 protein sequence.
 XX KW Human; HCDR-1; HCDR-2; HCDR-3; human cell division regulator; apoptosis;
 KW inflammation; cell proliferation disorder; adenocarcinoma; AIDS.
 XX OS Homo sapiens.
 XX PN US5871973-A.
 XX PD 16-FEB-1999.
 XX PF 15-OCT-1997; 97US-0951148.
 XX PR 15-OCT-1997; 97US-0951148.
 XX PA (INCY-) INCYTE PHARM INC.
 XX PI Bandman O, Corley NC, Hillman JL, Lal P, Shah P;
 XX DR WPI; 1999-166646/14.
 XX DR N-PSDB; AAX01577.

PT New polynucleotides encoding human cell division regulators (HCDR)
 PT useful for diagnosing, preventing and treating inflammation and
 PT disorders associated with cell proliferation and apoptosis
 XX PS Claim 1; Fig 1; 59pp; English.
 XX CC This sequence is the human cell division regulator-1 (HCDR-1)
 CC protein of the invention. Polynucleotides complementary to the HCDR-1
 CC coding sequence can be used as probes to detect the DNA in a sample. The
 CC polynucleotide sequences encoding HCDR may be used to prevent/treat
 CC inflammation and disorders associated with cell proliferation and
 CC apoptosis and in assays that detect activation of cancers.
 CC Polynucleotides encoding HCDR may be used for the diagnosis of conditions
 CC associated with expression of HCDR, including disorders associated with
 CC cell proliferation/apoptosis e.g. adenocarcinoma and AIDS. The
 CC polynucleotides may also be used in Southern or Northern analysis, dot
 CC blot, or other membrane based technologies; in PCR technologies; or in
 CC dipstick, pin, or ELISA assays or microarrays utilising fluids or tissues
 CC from patient biopsies to detect altered HCDR expression.
 XX SQ Sequence 478 AA;
 Query Match 100.0%; Score 27; DB 20; Length 478;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVD 5
 DB 190 KHAVD 194
 |||||
 RESULT 23
 AAB23259
 ID AAB23259 standard; Protein; 478 AA.
 XX AC AAB23259;
 XX DT 29-JAN-2001 (first entry)
 XX DE Human cell division regulator HCDR-1.
 XX KW Human cell division regulator; HCDR-1; spleen cell; HCDR-3;
 KW apoptotic disorder; AIDS; neurodegenerative disease; Alzheimer's disease;
 KW ischaemic injury; myelodysplastic syndrome; toxin-induced disease;
 KW viral infection; osteoporosis; inflammation; allergy; asthma; cancer;
 KW bacterial infection; fungal infection; parasitic infection; trauma;
 KW cellular proliferation; gene expression quantitation; drug screening;
 KW polymorphism detection; gene therapy.
 XX OS Homo sapiens.
 XX PN US6121019-A.
 XX PD 19-SEP-2000.
 XX PF 23-MAR-1999; 99US-0274570.
 XX PR 15-OCT-1997; 97US-0951148.
 XX PR 01-OCT-1998; 98US-0165234.
 XX PA (INCY-) INCYTE PHARM INC.
 XX PI Lal P, Shah P, Corley NC, Hillman JL, Bandman O;
 XX DR WPI; 2000-593713/56.
 XX DR N-PSDB; AAX97585.
 XX PT New human cell division regulator polypeptides and polynucleotides
 PT encoding them, useful for diagnosing, treating and preventing
 PT inflammation, cancer, and diseases associated with apoptosis
 XX PS Example VIII; Fig 1A-E; 59pp; English.
 XX

CC The invention relates to the human cell division regulator HCDR-3
CC (AAB23261) and to cDNA encoding it (AA097595), which was initially
CC isolated from an aortic endothelial cell line. The invention also relates
CC to expression constructs and host cells comprising HCDR-3 nucleotides,
CC methods of detecting HCDR-3 nucleotides, and the recombinant production
CC of HCDR-3. HCDR proteins, including HCDR-3 but also the related proteins
CC HCDR-1 (AAB23259) and HCDR-2 (AAB23260), and HCDR-encoding nucleotides
CC are useful in the diagnosis, prevention and treatment of disorders
CC associated with an increase in apoptosis such as AIDS and other
CC infectious or genetic immunodeficiencies, neurodegenerative diseases
CC (e.g., Alzheimer's disease), myelodysplastic syndromes, ischaemic
CC injuries, toxin-induced diseases (e.g., liver cirrhosis), wasting
CC diseases (e.g., cachexia), viral infections (e.g., hepatitis B and C), and
CC osteoporosis. HCDR proteins and nucleotides may also be used in the
CC treatment and prevention of inflammation (e.g., allergies, asthma),
CC disorders associated with cell proliferation (e.g., adenocarcinoma,
CC leukaemia, and other cancers), complications of cancer, haemodialysis,
CC extracorporeal circulation, viral, bacterial, fungal, parasitic,
CC protozoal and helminthic infections, and trauma. They may be further used
CC as hybridisation probes to detect and quantitate gene expression in
CC biopsied tissues in which expression of HCDRs may be correlated with
CC disease, and in assays to detect activation of induction of various
CC cancers. HCDRs may be added to cells to stimulate cell proliferation for
CC use in heterologous or autologous transplantation, or to fight infection
CC or cancer, or to correct a genetic defect which causes disease, e.g.,
CC sickle cell anaemia or beta-thalassaemia. The HCDR nucleotides may be
CC used to detect polymorphisms between normal, carrier or affected
CC individuals, as well as for screening libraries of compounds in drug
CC screening techniques. The present sequence represents HCDR-1, cDNA
CC encoding which was initially identified in a spleen cell cDNA library.

XX SQ Sequence 478 AA;

Query Match 100.0%; Score 27; DB 21; Length 478;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5

Db 190 KHAVD 194

RESULT 24

AAB12885

ID AAB12885 standard; Protein; 478 AA.

XX AC AAB12885;

XX DT 02-NOV-2000 (first entry)

XX DE Human beta-bradeion protein sequence.

XX KW Human; beta-bradeion; interleukin receptor; brain; neuronal cell death;
XX KW colon cancer; skin cancer; neuronal defect; aging; neuron necrosis;
XX KW cerebral ischaemia; epilepsy.

XX OS Homo sapiens.

XX FN JP2000139470-A.

XX PD 23-MAY-2000.

XX PF 16-NOV-1998; 98JP-0325380.

XX PR 16-NOV-1998; 98JP-0325380.

XX PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

XX WPI; 2000-415562/36.

DR N-PSDB; AAA62920.

XX PT Human-derived bradeion protein, its DNA and their applications

XX

PS Claim 4; Page 13-14; 25pp; Japanese.

XX This sequence represents a human beta-bradeion protein. The invention
CC relates to human derived bradeion proteins and the DNA sequences encoding
CC them. Bradeion is a membrane protein with a structure characteristic of
CC an interleukin receptor. Bradeion is highly expressed in the human adult
CC brain and at lower levels in the heart. Excessive expression of bradeion
CC in a human undifferentiated neuron culture causes cell death. Bradeion is
CC present in the cytoplasm during the lead up to cell death. The bradeion
CC gene is specifically expressed in human colon and skin cancer cell
CC strains. The bradeion protein and DNA sequences are useful in monitoring
CC neuronal defects caused by aging, and can also be used in the elucidation
CC of neuron necrosis caused by cerebral ischaemia. The DNA and protein
CC sequences may be of use when investigating epilepsy.

XX SQ Sequence 478 AA;

Query Match 100.0%; Score 27; DB 21; Length 478;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5

Db 190 KHAVD 194

RESULT 25

AAG78669

ID AAG78669 standard; Protein; 478 AA.

XX AC AAG78669;

XX DT 04-DEC-2001 (first entry)

XX DE Human bradeion protein #2.

XX KW Human; bradeion protein; cancer detection; interleukin receptor;
XX KW neurone; aging; neurone necrosis; cerebral ischaemia; epilepsy.

XX OS Homo sapiens.

XX FN JP2001161384-A.

XX PD 19-JUN-2001.

XX PF 16-NOV-1998; 2000JP-0308650.

XX PR 16-NOV-1998; 98JP-0325380.

XX PA (KEIZ-) KEIZAI SANGYOSHO SANGYO GIJUTSU SOGO KEN.

XX WPI; 2001-592468/67.

DR N-PSDB; AAI64992.

XX PT Detection of cancers, involves using DNA encoding Bradeion protein, or
XX PT antibodies specific for the protein

XX PS Claim 1; Page 13-14; 18pp; Japanese.

XX The present invention describes a method of detecting cancer, using an
CC antibody specific to bradeion protein or its coding sequence. The protein
CC has a structure similar to that of an interleukin receptor, and is highly
CC expressed in the human brain. In addition, the sequences can be used to
CC monitor the falling-off of neurones with increasing age and to elucidate
CC neurone necrosis which occurs in cerebral ischaemia and epilepsy
CC indigitation. The present sequence is a version of the bradeion protein.

XX SQ Sequence 478 AA;

Query Match 100.0%; Score 27; DB 22; Length 478;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KHAVD 5
|||
Db 190 KHAVD 194

Search completed: June 20, 2003, 20:51:52
Job time : 33.9318 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 10.5682 Seconds
(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-12

Perfect score: 27

Sequence: 1 KHAVD 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS.COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	5	3	US-08-893-534A-20
2	27	100.0	5	4	US-08-996-679-20
3	27	100.0	5	4	US-09-115-395-53
4	27	100.0	5	4	US-09-507-102-20
5	27	100.0	5	4	US-09-250-059-12
6	27	100.0	5	4	US-09-248-074-12
7	27	100.0	5	4	US-09-357-717-12
8	27	100.0	5	4	US-09-458-870-12
9	27	100.0	478	2	US-08-951-148-1
10	27	100.0	478	2	US-08-951-148-7
11	27	100.0	478	2	US-09-165-234-1
12	27	100.0	478	2	US-09-165-234-7
13	27	100.0	478	3	US-09-274-570-1
14	27	100.0	478	3	US-09-274-570-7
15	27	100.0	478	4	US-09-440-936-4
16	27	100.0	605	4	US-09-440-936-2
17	26	96.3	295	4	US-09-134-001C-4242
18	26	96.3	492	4	US-08-984-618-2
19	26	96.3	494	3	US-08-988-251-2
20	26	96.3	494	3	US-09-386-048-2
21	26	96.3	4150	4	US-09-428-517-2
22	24	88.9	313	1	US-08-302-449-2
23	24	88.9	313	5	PTC-US94-07430-2
24	24	88.9	319	1	US-07-745-206A-17
25	24	88.9	319	2	US-08-311-363-17
26	24	88.9	336	4	US-09-134-001C-5297
27	24	88.9	341	4	US-09-134-001C-5235

28	24	88.9	464	4	US-08-975-762-69	Sequence 69, Appl
29	24	88.9	464	4	US-09-295-028-69	Sequence 69, Appl
30	24	88.9	464	4	US-03-106-582-69	Sequence 69, Appl
31	24	88.9	562	4	US-09-227-717-2	Sequence 2, Appl
32	24	88.9	562	4	US-09-173-151A-22	Sequence 22, Appl
33	24	88.9	669	4	US-09-342-647-18	Sequence 18, Appl
34	24	88.9	854	2	US-08-928-692-17	Sequence 17, Appl
35	24	88.9	854	4	US-09-339-972-17	Sequence 17, Appl
36	24	88.9	1075	1	US-07-623-033-2	Sequence 2, Appl
37	24	88.9	1182	4	US-09-041-886-21	Sequence 21, Appl
38	24	88.9	1754	1	US-07-745-206A-13	Sequence 13, Appl
39	24	88.9	1754	2	US-08-311-363-13	Sequence 13, Appl
40	24	88.9	2237	1	US-08-455-543A-48	Sequence 48, Appl
41	24	88.9	2237	2	US-08-223-305C-48	Sequence 48, Appl
42	24	88.9	2237	4	US-08-268-163-8	Sequence 8, Appl
43	24	88.9	2265	2	US-08-149-097D-36	Sequence 36, Appl
44	24	88.9	2336	4	US-08-268-163-10	Sequence 10, Appl
45	24	88.9	2337	3	US-08-713-118-2	Sequence 2, Appl
46	24	88.9	2337	4	US-09-452-007-2	Sequence 2, Appl
47	24	88.9	2339	1	US-08-455-543A-47	Sequence 47, Appl
48	24	88.9	2339	2	US-08-223-305C-47	Sequence 47, Appl
49	24	88.9	2339	4	US-09-268-163-6	Sequence 6, Appl
50	24	88.9	2343	4	US-08-149-097D-35	Sequence 35, Appl
51	24	88.9	2509	2	US-08-893-534A-21	Sequence 21, Appl
52	23	85.2	5	3	US-08-996-679-21	Sequence 21, Appl
53	23	85.2	5	4	US-08-996-679-56	Sequence 56, Appl
54	23	85.2	5	4	US-09-115-395-18	Sequence 18, Appl
55	23	85.2	5	4	US-09-115-395-54	Sequence 54, Appl
56	23	85.2	5	4	US-09-507-102-21	Sequence 21, Appl
57	23	85.2	5	4	US-09-250-059-13	Sequence 13, Appl
58	23	85.2	5	4	US-09-250-059-16	Sequence 16, Appl
59	23	85.2	5	4	US-09-248-074-13	Sequence 13, Appl
60	23	85.2	5	4	US-09-248-074-16	Sequence 16, Appl
61	23	85.2	5	4	US-09-357-717-14	Sequence 14, Appl
62	23	85.2	5	4	US-09-458-870-13	Sequence 13, Appl
63	23	85.2	5	4	US-09-458-870-16	Sequence 16, Appl
64	23	85.2	5	4	US-07-861-458C-139	Sequence 139, Appl
65	23	85.2	20	3	US-08-817-926-3	Sequence 3, Appl
66	23	85.2	76	1	US-08-393-985-19	Sequence 19, Appl
67	23	85.2	114	4	US-09-325-932A-196	Sequence 196, Appl
68	23	85.2	213	3	US-08-523-373-22	Sequence 22, Appl
69	23	85.2	213	3	US-08-523-373-23	Sequence 23, Appl
70	23	85.2	214	3	US-08-523-373-24	Sequence 24, Appl
71	23	85.2	215	3	US-08-523-373-24	Sequence 24, Appl
72	23	85.2	263	1	US-08-393-985-23	Sequence 23, Appl
73	23	85.2	266	4	US-09-631-828A-4	Sequence 4, Appl
74	23	85.2	283	4	US-09-134-001C-5652	Sequence 5652, Ap
75	23	85.2	285	1	US-08-393-985-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-893-534A-20
; Sequence 20, Application US/08893534A
; Patent No. 6031072
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893.534A
FILING DATE: 11-JUL-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-893-534A-20

Query Match 100.0%; Score 27; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 1 KHAVD 5

RESULT 2

US-08-996-679-20
Sequence 20, Application US/08996679
Patent No. 6169071
GENERAL INFORMATION:

APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996.679
FILING DATE: 23-DEC-1997

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-20

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVD 5
Db 1 KHAVD 5

RESULT 3

US-09-115-395-53
Sequence 53, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115.395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996.679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893.534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 53
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis
US-09-115-395-53

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 1 KHAVD 5

RESULT 4

US-09-507-102-20
Sequence 20, Application US/09507102
Patent No. 6328352
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED IP LAW GROUP PLLC
STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507.102
FILING DATE: 17-Feb-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/893,534
FILING DATE: 11-JUL-1997
APPLICATION NUMBER: US 60/021,612
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Christiansen, William T.
REGISTRATION NUMBER: 44,614
REFERENCE/DOCKET NUMBER: 100086.401C10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-507-102-20

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 1 KHAVD 5

RESULT 5

US-09-250-059-12
Sequence 12, Application US/09250059
Patent No. 633307
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250,059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-250-059-12

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 1 KHAVD 5

RESULT 6

US-09-248-074-12
Sequence 12, Application US/09248074
Patent No. 6346512
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.

APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
FILE REFERENCE: 100086.401C5
CURRENT APPLICATION NUMBER: US/09/248,074
CURRENT FILING DATE: 1999-02-10
NUMBER OF SEQ ID NOS: 81
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-248-074-12

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 1 KHAVD 5

RESULT 7

US-09-357-717-12
Sequence 12, Application US/09357717
Patent No. 6417325
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
APPLICANT: Farookhi, Riaz
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
FILE REFERENCE: 100086.401C7
CURRENT APPLICATION NUMBER: US/09/357,717
CURRENT FILING DATE: 1999-07-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-357-717-12

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 1 KHAVD 5

RESULT 8

US-09-458-870-12
Sequence 12, Application US/09458870
Patent No. 6465427

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
APPLICANT: Farookhi, Riaz
APPLICANT: Ali, Ammar
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
FILE REFERENCE: 100086.401C8
CURRENT APPLICATION NUMBER: US/09/458,870
CURRENT FILING DATE: 1999-12-10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-458-870-12

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 1 KHAVD 5

RESULT 9
US-08-951-148-1
Sequence 1, Application US/08951148
Patent No. 5871973
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: CELL DIVISION REGULATORS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,148
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0407 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SPLNFT01
CLONE: 26459
US-08-951-148-1

Query Match 100.0%; Score 27; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 190 KHAVD 194

RESULT 10
US-08-951-148-7
Sequence 7, Application US/08951148
Patent No. 5871973
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: CELL DIVISION REGULATORS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,148
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0407 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 51203
US-08-951-148-7

Query Match 100.0%; Score 27; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
Db 190 KHAVD 194

RESULT 11
US-09-165-234-1
; Sequence 1, Application US/09165234
; Patent No. 5928899
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: CELL DIVISION REGULATORS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,234
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/951,148
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0407 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNFT01
; CLONE: 26459
US-09-165-234-1

Query Match 100.0%; Score 27; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
Db 190 KHAVD 194

RESULT 12
US-09-165-234-7
; Sequence 7, Application US/09165234
; Patent No. 5928899
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: CELL DIVISION REGULATORS

; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,234
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/951,148
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0407 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 51203
US-09-165-234-7

Query Match 100.0%; Score 27; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
Db 190 KHAVD 194

RESULT 13
US-09-274-570-1
; Sequence 1, Application US/09274570
; Patent No. 6121019
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: CELL DIVISION REGULATORS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/274,570

;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/951,148
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0407 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-855-0555
;; TELEFAX: 650-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 478 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: SPLNPFZT01
;; CLONE: 26459
US-09-274-570-1

Query Match 100.0%; Score 27; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVD 5
DB 190 KHAVD 194

RESULT 14
US-09-274-570-7
; Sequence 7, Application US/09274570
; Patent No. 6121019
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: CELL DIVISION REGULATORS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FabSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/274,570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/951,148
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0407 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 478 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GenBank
;; CLONE: 51203
US-09-274-570-7

Query Match 100.0%; Score 27; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 190 KHAVD 194

RESULT 15
US-09-440-936-4
; Sequence 4, Application US/09440936
; Patent No. 6423504
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Manami
; APPLICANT: TANAKA, Tomoo
; TITLE OF INVENTION: HUMAN-DERIVED BRADETON PROTEINS, DNA CODING FOR THE
; FILE REFERENCE: 081356/0138
; CURRENT APPLICATION NUMBER: US/09/440,936
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: JP 325380/1998
; EARLIER FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-440-936-4

Query Match 100.0%; Score 27; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 190 KHAVD 194

RESULT 16
US-09-440-936-2
; Sequence 2, Application US/09440936
; Patent No. 6423504
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Manami
; APPLICANT: TANAKA, Tomoo
; TITLE OF INVENTION: HUMAN-DERIVED BRADETON PROTEINS, DNA CODING FOR THE
; FILE REFERENCE: 081356/0138
; CURRENT APPLICATION NUMBER: US/09/440,936
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: JP 325380/1998
; EARLIER FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-440-936-2

Query Match 100.0%; Score 27; DB 4; Length 605;
Best Local Similarity 100.0%; Pred. No. 78;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
|||||
Db 313 KHAVD 317

RESULT 17

US-09-134-001C-4242
; Sequence 4242, Application US/09134001C
; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4242

; LENGTH: 295

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4242

Query Match 96.3%; Score 26; DB 4; Length 295;

Best Local Similarity 80.0%; Pred. No. 61;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
|||||
Db 161 KHAID 165

RESULT 18

US-08-984-618-2

; Sequence 2, Application US/08984618

; Patent No. 6251647

; GENERAL INFORMATION:

; APPLICANT: de Lencastre, Herminia

; TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF

; TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/984,618

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-213 CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-487-5800

; TELEFAX: 201-343-1684

; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 492 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Staphylococcus aureus

US-08-984-618-2

Query Match

Best Local Similarity 96.3%; Score 26; DB 4; Length 492;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5

|||||

Db 439 KHAID 443

RESULT 19

US-08-988-251-2

; Sequence 2, Application US/08988251

; Patent No. 6013497

; GENERAL INFORMATION:

; APPLICANT: Wallis, Nicola G.

; TITLE OF INVENTION: Mure

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dechert Price & Rhoads

; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre

; CITY: Philadelphia

; STATE: PA

; COUNTRY: US

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/988,251

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Dickinson, Todd Q

; REGISTRATION NUMBER: 28,354

; REFERENCE/DOCKET NUMBER: GM10094

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-994-2252

; TELEFAX: 215-994-2222

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 494 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-988-251-2

Query Match

Best Local Similarity 96.3%; Score 26; DB 3; Length 494;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5

|||||

Db 440 KHAID 444

RESULT 20

US-09-386-048-2
; Sequence 2, Application US/09386048
; Patent No. 6103507

; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola
; APPLICANT: Shilling, Lisa
; APPLICANT: Zabeierowski, Susan
; TITLE OF INVENTION: Mure
; FILE REFERENCE: GM10094-D1
; CURRENT APPLICATION NUMBER: US/09/386,048
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: US 08/988,251
; EARLIER FILING DATE: 1997-12-10
; EARLIER APPLICATION NUMBER: US 60/059,956
; EARLIER FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-386-048-2

Query Match 96.3%; Score 26; DB 3; Length 494;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KHAVD 5
|||:|
Db 440 KHAID 444

RESULT 21

US-09-428-517-2
; Sequence 2, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2

Query Match 96.3%; Score 26; DB 4; Length 4150;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KHAVD 5
|||:|
Db 2541 KHAID 2545

RESULT 22

US-08-302-449-2
; Sequence 2, Application US/08302449
; Patent No. 5679635

; GENERAL INFORMATION:
; APPLICANT: Matalon, Reuben
; APPLICANT: Kaul, Rajinder
; APPLICANT: Cao, Guang Ping
; APPLICANT: Balamurugan, Kuppareddi
; APPLICANT: Michals-Matalon, Kimberlee
; TITLE OF INVENTION: Aspartocyclase Gene, Protein, and
; TITLE OF INVENTION: Methods of Screening for Mutations Associated with Canav
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Boulevard, Suite 1400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,449
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/128,020
; FILING DATE: 29-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07430
; FILING DATE: 05-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: SHUTT 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; TELEX: 64191
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 83
; OTHER INFORMATION: /note= "Phosphorylation site"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 105
; OTHER INFORMATION: /note= "Phosphorylation site"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 108
; OTHER INFORMATION: /note= "Phosphorylation site"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 146
; OTHER INFORMATION: /note= "Phosphorylation site"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 264
; OTHER INFORMATION: /note= "Phosphorylation site"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 117
; OTHER INFORMATION: /note= "Potential N-glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 18..24

OTHER INFORMATION: /note= "Consensus sequence
OTHER INFORMATION: predicted to be involved in catalysis"
FEATURE:
NAME/KEY: Active-site
LOCATION: 275..278
OTHER INFORMATION: /note= "Consensus sequence
OTHER INFORMATION: predicted to be involved in catalysis"
FEATURE:
NAME/KEY: Active-site
LOCATION: 283..289
OTHER INFORMATION: /note= "Consensus sequence
OTHER INFORMATION: predicted to be involved in catalysis"
US-08-302-449-2

Query Match 88.9%; Score 24; DB 1; Length 313;
Best Local Similarity 80.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0;

OY 1 KHAVD 5
|||:|
DB 200 KHALD 204

RESULT 23
PCT-US94-07430-2
Sequence 2, Application PC/TUS9407430
GENERAL INFORMATION:
APPLICANT: Matalon, Reuben
APPLICANT: Kaul, Rajinder
APPLICANT: Cao, Guang Ping
APPLICANT: Balamurugan, Kuppareddi
APPLICANT: Michals-Matalon, Kimberlee
TITLE OF INVENTION: Aspartoacylase Gene, Protein, and
TITLE OF INVENTION: Methods of Screening for Mutations Associated with Canavan
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Boulevard, Suite 1400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07430
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/128,020
FILING DATE: 29-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hamlet-King, Diana
REGISTRATION NUMBER: 33,302
REFERENCE/DOCKET NUMBER: SHUTT 1PO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
TELEX: 64191

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 83
OTHER INFORMATION: /note= "Phosphorylation site"
FEATURE:

NAME/KEY: Modified-site
LOCATION: 105
OTHER INFORMATION: /note= "Phosphorylation site"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 108
OTHER INFORMATION: /note= "Phosphorylation site"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 146
OTHER INFORMATION: /note= "Phosphorylation site"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 264
OTHER INFORMATION: /note= "phosphorylation site"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 117
OTHER INFORMATION: /note= "Potential N-glycosylation
OTHER INFORMATION: site"
FEATURE:
NAME/KEY: Active-site
LOCATION: 18..24
OTHER INFORMATION: /note= "Consensus sequence
OTHER INFORMATION: predicted to be involved in catalysis"
FEATURE:
NAME/KEY: Active-site
LOCATION: 275..278
OTHER INFORMATION: /note= "Consensus sequence
OTHER INFORMATION: predicted to be involved in catalysis"
FEATURE:
NAME/KEY: Active-site
LOCATION: 283..289
OTHER INFORMATION: /note= "Consensus sequence
OTHER INFORMATION: predicted to be involved in catalysis"
PCT-US94-07430-2

Query Match 88.9%; Score 24; DB 5; Length 313;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVD 5
|||:|
DB 200 KHALD 204

RESULT 24
US-07-745-206A-17
Sequence 17, Application US/07745206A
Patent No. 5429321
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: Human Calcium Channel Compositions and
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A

FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 51504
TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 319 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-745-206A-17

Query Match 88.9%; Score 24; DB 1; Length 319;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 134 KHSVD 138

RESULT 25

US-08-311-363-17
Sequence 17, Application US/08311363
Patent No. 5876958
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: Human Calcium Channel Compositions and
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,363
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-51506
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 319 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-311-363-17

Query Match 88.9%; Score 24; DB 2; Length 319;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVD 5
Db 134 KHSVD 138
Search completed: June 20, 2003, 21:00:29
Job time: 12.5682 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:52:19 ; Search time 12.8409 Seconds
(without alignments)
42.134 Million cell updates/sec

Title: US-10-105-008-12

Perfect score: 27

Sequence: 1 KHAVD 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	5	9	US-09-769-145-12
2	27	100.0	5	9	US-10-058-821-12
3	27	100.0	5	10	US-09-234-395-309
4	27	100.0	5	10	US-09-305-928-309
5	27	100.0	5	12	US-10-006-982-20
6	27	100.0	32	9	US-10-174-410-192
7	27	100.0	159	9	US-09-738-626-5261
8	27	100.0	478	9	US-10-190-555-4
9	27	100.0	605	9	US-10-190-555-2
10	26	96.3	429	10	US-09-815-242-12891
11	26	96.3	429	10	US-09-815-242-13109
12	26	96.3	492	10	US-09-815-242-5808
13	26	96.3	493	9	US-10-084-205-24
14	26	96.3	433	10	US-09-925-637-24
15	26	96.3	4150	9	US-09-808-880-2
16	24	88.9	129	9	US-10-079-623-353
17	24	88.9	159	9	US-09-746-783-130
18	24	88.9	171	10	US-09-925-301-1143
19	24	88.9	371	9	US-09-992-598-171

ALIGNMENTS

RESULT 1:
US-09-769-145-12

; Sequence 12, Application US/09769145
; Patent No. US20020168761A1

; GENERAL INFORMATION:

; APPLICANT: Gour, Barbara J.

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Ali, Annmar

; APPLICANT: Ni, Feng

; APPLICANT: Chen, Zhigang

; APPLICANT: Michaud, Stephanie

; APPLICANT: Wang, Shoameng

20	24	88.9	371	9	US-09-989-293A-171	Sequence 171, App
21	24	88.9	371	9	US-09-989-735-171	Sequence 171, App
22	24	88.9	371	9	US-09-990-444-171	Sequence 171, App
23	24	88.9	371	9	US-09-989-730-171	Sequence 171, App
24	24	88.9	371	9	US-09-990-436-171	Sequence 171, App
25	24	88.9	371	9	US-09-991-181-171	Sequence 171, App
26	24	88.9	371	9	US-09-993-687-171	Sequence 171, App
27	24	88.9	371	9	US-09-989-734-171	Sequence 171, App
28	24	88.9	371	9	US-09-997-653-171	Sequence 171, App
29	24	88.9	371	9	US-10-174-590-146	Sequence 146, App
30	24	88.9	371	9	US-10-176-758-146	Sequence 146, App
31	24	88.9	371	9	US-10-175-737-146	Sequence 146, App
32	24	88.9	371	9	US-09-993-667-171	Sequence 171, App
33	24	88.9	371	9	US-10-173-706-146	Sequence 146, App
34	24	88.9	371	9	US-10-175-738-146	Sequence 146, App
35	24	88.9	371	9	US-10-175-752-146	Sequence 146, App
36	24	88.9	371	9	US-10-176-482-146	Sequence 146, App
37	24	88.9	371	9	US-10-176-757-146	Sequence 146, App
38	24	88.9	371	9	US-10-176-913-146	Sequence 146, App
39	24	88.9	371	9	US-10-180-553-146	Sequence 146, App
40	24	88.9	371	9	US-10-180-557-146	Sequence 146, App
41	24	88.9	371	9	US-09-990-438-171	Sequence 171, App
42	24	88.9	371	9	US-09-990-562-171	Sequence 171, App
43	24	88.9	371	9	US-09-997-428-171	Sequence 171, App
44	24	88.9	371	9	US-09-997-666-171	Sequence 171, App
45	24	88.9	371	9	US-10-173-700-146	Sequence 146, App
46	24	88.9	371	9	US-10-174-572-146	Sequence 146, App
47	24	88.9	371	9	US-10-174-573-146	Sequence 146, App
48	24	88.9	371	9	US-10-174-582-146	Sequence 146, App
49	24	88.9	371	9	US-10-175-739-146	Sequence 146, App
50	24	88.9	371	9	US-10-175-740-146	Sequence 146, App
51	24	88.9	371	9	US-10-175-743-146	Sequence 146, App
52	24	88.9	371	9	US-10-176-488-146	Sequence 146, App
53	24	88.9	371	9	US-10-176-492-146	Sequence 146, App
54	24	88.9	371	9	US-10-176-747-146	Sequence 146, App
55	24	88.9	371	9	US-10-176-747-146	Sequence 146, App
56	24	88.9	371	9	US-10-176-750-146	Sequence 146, App
57	24	88.9	371	9	US-10-176-985-146	Sequence 146, App
58	24	88.9	371	9	US-10-176-987-146	Sequence 146, App
59	24	88.9	371	9	US-10-176-991-146	Sequence 146, App
60	24	88.9	371	9	US-10-176-992-146	Sequence 146, App
61	24	88.9	371	9	US-10-176-993-146	Sequence 146, App
62	24	88.9	371	9	US-10-184-658-146	Sequence 144, App
63	24	88.9	371	9	US-10-227-884-144	Sequence 171, App
64	24	88.9	371	9	US-09-990-711-171	Sequence 146, App
65	24	88.9	371	9	US-10-173-693-146	Sequence 146, App
66	24	88.9	371	9	US-10-173-697-146	Sequence 146, App
67	24	88.9	371	9	US-10-173-705-146	Sequence 146, App
68	24	88.9	371	9	US-10-174-576-146	Sequence 146, App
69	24	88.9	371	9	US-10-174-585-146	Sequence 146, App
70	24	88.9	371	9	US-10-174-586-146	Sequence 146, App
71	24	88.9	371	9	US-10-176-481-146	Sequence 146, App
72	24	88.9	371	9	US-10-176-485-146	Sequence 146, App
73	24	88.9	371	9	US-10-176-487-146	Sequence 146, App
74	24	88.9	371	9	US-10-176-487-146	Sequence 146, App
75	24	88.9	371	9	US-10-176-493-146	Sequence 146, App

APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 10086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-12

Query Match 100.0%; Score 27; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVD 5
Db 1 KHAVD 5

RESULT 2
US-10-058-821-12
Sequence 12, Application US/10058821
Publication No. US20030087811A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
FILE REFERENCE: 10086.401C12
CURRENT APPLICATION NUMBER: US/10/058,821
CURRENT FILING DATE: 2002-01-29
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-10-058-821-12

Query Match 100.0%; Score 27; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVD 5
Db 1 KHAVD 5

RESULT 3
US-09-234-395-309
Sequence 309, Application US/09234395
Patent No. US20020123044A1

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 10086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 309
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-234-395-309

Query Match 100.0%; Score 27; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVD 5
Db 1 KHAVD 5

RESULT 4
US-09-305-928-309
Sequence 309, Application US/09305928
Patent No. US20020146687A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 10086.407C4
CURRENT APPLICATION NUMBER: US/09/305,928
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 309
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-305-928-309

Query Match 100.0%; Score 27; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVD 5
Db 1 KHAVD 5

RESULT 5
US-10-006-982-20
Sequence 20, Application US/10006982
Patent No. US20020151475A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED IP LAW GROUP PLLC
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/006,982
FILING DATE: 04-Dec-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Christiansen, William T.
REGISTRATION NUMBER: 44,614
REFERENCE/DOCKET NUMBER: 100086.401C11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-006-982-20

Query Match 100.0%; Score 27; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 1 KHAVD 5

RESULT 6
US-10-174-410-192
; Sequence 192, Application US/10174410
; Publication No. US20030073134A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; FILE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs
; FILE REFERENCE: 524982000300
; CURRENT APPLICATION NUMBER: US/10/174,410
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/299,058
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-174-410-192

Query Match 100.0%; Score 27; DB 9; Length 32;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 8 KHAVD 12

RESULT 7
US-09-738-626-5261
; Sequence 5261, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5261
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5261

Query Match 100.0%; Score 27; DB 9; Length 159;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 36 KHAVD 40

RESULT 8
US-10-190-555-4
; Sequence 4, Application US/10190555
; Publication No. US20030099970A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Manami
; APPLICANT: TANAKA, Tomoo
; TITLE OF INVENTION: HUMAN-DERIVED BRADEN PROTEINS, AND USES THEREOF
; FILE OF INVENTION: 081356/0138
; FILE REFERENCE: 081356/0138
; CURRENT APPLICATION NUMBER: US/10/190,555
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/440,936
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: JP 325380/1998
; PRIOR FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-555-4

Query Match 100.0%; Score 27; DB 9; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
| | | | |
Db 190 KHAVD 194

RESULT 9

US-10-190-555-2

; Sequence 2, Application US/10190555

; Publication No. US2003009970A1

; GENERAL INFORMATION:

; APPLICANT: TANAKA, Manami

; APPLICANT: TANAKA, Tomoo

; TITLE OF INVENTION: HUMAN-DERIVED BRADEION PROTEINS, DNA CODING FOR THE

; FILE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: 081356/0138

; CURRENT APPLICATION NUMBER: US/10/190,555

; PRIOR FILING DATE: 2002-07-09

; PRIOR APPLICATION NUMBER: US/09/440,936

; PRIOR FILING DATE: 1999-11-16

; PRIOR APPLICATION NUMBER: JP 325380/1998

; PRIOR FILING DATE: 1998-11-16

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 605

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-190-555-2

Query Match

Best Local Similarity 100.0%; Score 27; DB 9; Length 605;

; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
| | | | |
Db 313 KHAVD 317

RESULT 10

US-09-815-242-12891

; Sequence 12891, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 12891

; LENGTH: 429

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-12891

Query Match

Best Local Similarity 96.3%; Score 26; DB 10; Length 429;

; Pred. No. 2.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
| | | | |
Db 376 KHAVD 380

RESULT 11

US-09-815-242-13109

; Sequence 13109, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 13109

; LENGTH: 429

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-13109

Query Match

Best Local Similarity 96.3%; Score 26; DB 10; Length 429;

; Pred. No. 2.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
| | | | |
Db 376 KHAVD 380

RESULT 12

US-09-815-242-5808

; Sequence 5808, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5808
LENGTH: 492
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5808

Query Match 96.3%; Score 26; DB 10; Length 492;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 439 KHAID 443

RESULT 13
US-10-084-205-24
Sequence 24, Application US/10084205
Publication No. US20030049648A1
GENERAL INFORMATION:
APPLICANT: Choi, Gil
FILE REFERENCE: PB515P1
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US/10/084,205
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: PCT/US00/23773
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 60/151,933
PRIOR FILING DATE: 1999-09-01
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 24
LENGTH: 493
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-084-205-24

Query Match 96.3%; Score 26; DB 9; Length 493;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 440 KHAID 444

RESULT 14

US-09-925-637-24
Sequence 24, Application US/09925637
Patent No. US2002010338A1
GENERAL INFORMATION:
APPLICANT: Choi
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
FILE REFERENCE: PB560
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: US/09/925,637
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 60/151,933
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: US 08/781,986
PRIOR FILING DATE: 1997-01-03
PRIOR APPLICATION NUMBER: US 08/956,171
PRIOR FILING DATE: 1997-10-20
PRIOR APPLICATION NUMBER: US 60/009,861
PRIOR FILING DATE: 1996-01-06
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 493
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-637-24

Query Match 96.3%; Score 26; DB 10; Length 493;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 440 KHAID 444

RESULT 15
US-09-808-880-2
Sequence 2, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/808,880
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-808-880-2

Query Match 96.3%; Score 26; DB 9; Length 4150;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 2541 KHAID 2545

RESULT 16

US-10-079-623-353
; Sequence 353, Application US/10079623
; Patent No. US20020169302A1
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Glenn, Matthew R.
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; TITLE OF INVENTION: Compositions isolated from bovine
; mammary gland and methods for their use.
; FILE REFERENCE: 11000.1046C3
; CURRENT APPLICATION NUMBER: US/10/079,623
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-623-353

Query Match 88.9%; Score 24; DB 9; Length 129;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
|||
Db 61 KHSVD 65

RESULT 17

US-09-746-783-130
; Sequence 130, Application US/09746783
; Publication No. US2003004935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; LaValle, Edward R.
; Racie, Lisa A.
; Treacy, Maurice
; Spaulding, Vikki
; Agostino, Michael J.
; Howes, Steven H.
; Fachtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/746,783

FILING DATE: 21-Dec-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Milasincic, Debra J.

REGISTRATION NUMBER: 46,931

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 130:

SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 130:
US-09-746-783-130

Query Match 88.9%; Score 24; DB 9; Length 159;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
|||
Db 56 RHAVD 60

RESULT 18

US-09-925-301-1143
; Sequence 1143, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1143
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1143

Query Match 88.9%; Score 24; DB 10; Length 171;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
|||
Db 53 KHSVD 57

RESULT 19

US-09-992-598-171
; Sequence 171, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2730PIC20
 CURRENT APPLICATION NUMBER: US/09/992,598
 CURRENT FILING DATE: 2001-11-14
 PRIOR APPLICATION NUMBER: 60/049787
 PRIOR FILING DATE: 1997-06-16
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/075945
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088734
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088742
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090676

PRIOR APPLICATION NUMBER: 60/090678
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090690
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090694
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090695
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090696
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090862
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091478
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091544
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091626
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091633
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091978
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/092182
 PRIOR FILING DATE: 1998-07-09

Query Match 88.9%; Score 24; DB 9; Length 371;
 Best Local Similarity 80.0%; Pred.No. 6.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KHAVD 5
 Db 56 RHAVD 60

RESULT 20

US-09-989-293A-171
 Sequence 171, Application US/09989293A
 Patent No. US20020177164A1
 GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi J.
 APPLICANT: Baker, Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Kljavin, Ivar J.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same

FILE REFERENCE: P2730PIC66
 CURRENT APPLICATION NUMBER: US/09/989,293A
 CURRENT FILING DATE: 2001-11-20
 PRIOR APPLICATION NUMBER: 60/049787
 PRIOR FILING DATE: 1997-06-16
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/075945
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088734
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088742
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 88.9%; Score 24; DB 9; Length 371;

Best Local Similarity 80.0%; Pred. No. 6.6e-02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 1;

Qy 1 KHAVD 5
Db 56 RHAVD 60
:||||

RESULT 21

US-09-989-735-171
; Sequence 171, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC61
; CURRENT APPLICATION NUMBER: US/09/989,735
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

1	PRIOR FILING DATE: 1998-06-16	
2	PRIOR APPLICATION NUMBER: 60/089514	
3	PRIOR FILING DATE: 1998-06-16	
4	PRIOR APPLICATION NUMBER: 60/089532	
5	PRIOR FILING DATE: 1998-06-17	
6	PRIOR APPLICATION NUMBER: 60/089538	
7	PRIOR FILING DATE: 1998-06-17	
8	PRIOR APPLICATION NUMBER: 60/089598	
9	PRIOR FILING DATE: 1998-06-17	
10	PRIOR APPLICATION NUMBER: 60/089599	
11	PRIOR FILING DATE: 1998-06-17	
12	PRIOR APPLICATION NUMBER: 60/089600	
13	PRIOR FILING DATE: 1998-06-17	
14	PRIOR APPLICATION NUMBER: 60/089653	
15	PRIOR FILING DATE: 1998-06-17	
16	PRIOR APPLICATION NUMBER: 60/089801	
17	PRIOR FILING DATE: 1998-06-18	
18	PRIOR APPLICATION NUMBER: 60/089907	
19	PRIOR FILING DATE: 1998-06-18	
20	PRIOR APPLICATION NUMBER: 60/089908	
21	PRIOR FILING DATE: 1998-06-18	
22	PRIOR APPLICATION NUMBER: 60/089947	
23	PRIOR FILING DATE: 1998-06-19	
24	PRIOR APPLICATION NUMBER: 60/089948	
25	PRIOR FILING DATE: 1998-06-19	
26	PRIOR APPLICATION NUMBER: 60/089952	
27	PRIOR FILING DATE: 1998-06-19	
28	PRIOR APPLICATION NUMBER: 60/090246	
29	PRIOR FILING DATE: 1998-06-22	
30	PRIOR APPLICATION NUMBER: 60/090252	
31	PRIOR FILING DATE: 1998-06-22	
32	PRIOR APPLICATION NUMBER: 60/090254	
33	PRIOR FILING DATE: 1998-06-22	
34	PRIOR APPLICATION NUMBER: 60/090349	
35	PRIOR FILING DATE: 1998-06-23	
36	PRIOR APPLICATION NUMBER: 60/090355	
37	PRIOR FILING DATE: 1998-06-23	
38	PRIOR APPLICATION NUMBER: 60/090429	
39	PRIOR FILING DATE: 1998-06-24	
40	PRIOR APPLICATION NUMBER: 60/090431	
41	PRIOR FILING DATE: 1998-06-24	
42	PRIOR APPLICATION NUMBER: 60/090435	
43	PRIOR FILING DATE: 1998-06-24	
44	PRIOR APPLICATION NUMBER: 60/090444	
45	PRIOR FILING DATE: 1998-06-24	
46	PRIOR APPLICATION NUMBER: 60/090445	
47	PRIOR FILING DATE: 1998-06-24	
48	PRIOR APPLICATION NUMBER: 60/090472	
49	PRIOR FILING DATE: 1998-06-24	
50	PRIOR APPLICATION NUMBER: 60/090535	
51	PRIOR FILING DATE: 1998-06-24	
52	PRIOR APPLICATION NUMBER: 60/090540	
53	PRIOR FILING DATE: 1998-06-24	
54	PRIOR APPLICATION NUMBER: 60/090542	
55	PRIOR FILING DATE: 1998-06-24	
56	PRIOR APPLICATION NUMBER: 60/090557	
57	PRIOR FILING DATE: 1998-06-24	
58	PRIOR APPLICATION NUMBER: 60/090676	
59	PRIOR FILING DATE: 1998-06-25	
60	PRIOR APPLICATION NUMBER: 60/090678	
61	PRIOR FILING DATE: 1998-06-25	
62	PRIOR APPLICATION NUMBER: 60/090690	
63	PRIOR FILING DATE: 1998-06-25	
64	PRIOR APPLICATION NUMBER: 60/090696	
65	PRIOR FILING DATE: 1998-06-25	
66	PRIOR APPLICATION NUMBER: 60/090862	
67	PRIOR FILING DATE: 1998-06-26	
68	PRIOR APPLICATION NUMBER: 60/090863	
69	PRIOR FILING DATE: 1998-06-26	

;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 88.9%; Score 24; DB 9; Length 371;

Best Local Similarity 80.0%; Pred. No. 6.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRAVD 5
Db 56 RHAVD 60

RESULT 22

US-09-990-444-171

; Sequence 171, Application US/09990444

; Publication No. US20020193300A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C19
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945

;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089598
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089599
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089600
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089653
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519

; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 88.9%; Score 24; DB 9; Length 371;
 Best Local Similarity 80.0%; Pred. No. 6.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
 Db 56 RHAVD 60

RESULT 23

US-09-989-730-171
 ; Sequence 171, Application US/09989730
 ; Publication No. US20020197674A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Deenoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2730P1C69
 ; CURRENT APPLICATION NUMBER: US/09/989,730
 ; CURRENT FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07

PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/098901
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089908
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089947
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089948
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089952
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090246
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090252
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090254
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090349
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090355
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090429
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090431
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090435
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090444
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090445
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090472
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090535
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090540
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090542
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090557
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090676
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090678
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090690
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090694
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090695
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090696
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090862
PRIOR FILING DATE:	1998-06-26
PRIOR APPLICATION NUMBER:	60/090863
PRIOR FILING DATE:	1998-06-26
PRIOR APPLICATION NUMBER:	60/091360
PRIOR FILING DATE:	1998-07-01
PRIOR APPLICATION NUMBER:	60/091478
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091544
PRIOR FILING DATE:	1998-07-01
PRIOR APPLICATION NUMBER:	60/091519
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091626
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091633
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091978
PRIOR FILING DATE:	1998-07-02

;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 88.9%; Score 24; DB 9; Length 371;
Best Local Similarity 80.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 56 RHAVD 60

RESULT 24

US-09-990-436-171
; Sequence 171, Application US/09990436
; Publication No. US20020198148A1

GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi J.
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan L.
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Kijavir, Ivar J.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2730PIC14
;; CURRENT APPLICATION NUMBER: US/09/990,436
;; PRIOR FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759

;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 88.9%; Score 24; DB 9; Length 371;
Best Local Similarity 80.0%; Pred. No. 6.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVD:5
Db 56 RHAVD 60
RESULT 25
US-09-991-181-171
; Sequence 171, Application US/09991181
; Publication No. US20020197615A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumaas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC53
; CURRENT APPLICATION NUMBER: US/09/991,181
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246

PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 88.9%; Score 24; DB 9; Length 371;
Best Local Similarity 80.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 56 RHAVD 60

us-10-105-008-12.rapb

Mon Jun 23 10:08:24 2003

Search completed: June 20, 2003, 21:02:32
Job time : 15.8409 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 12.7273 Seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-12

Perfect score: 27

Sequence: 1 KHAVD 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

PIR73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27	100.0	106	2	G82795	hypothetical prote
2	27	100.0	117	2	T02603	hypothetical prote
3	27	100.0	141	2	F64214	histidine triad pr
4	27	100.0	159	2	T05656	hypothetical prote
5	27	100.0	199	2	E89972	hypothetical prote
6	27	100.0	211	2	C84888	hypothetical prote
7	27	100.0	222	2	D64306	hypothetical prote
8	27	100.0	236	2	H71032	hypothetical prote
9	27	100.0	265	2	H84422	probable ribose 5-
10	27	100.0	382	2	F87464	conserved hypothet
11	27	100.0	427	2	A13490	membrane-bound lyt
12	27	100.0	431	2	G81179	oxidoreductase, pr
13	27	100.0	478	2	S16867	gene H5 protein -
14	27	100.0	481	2	A95278	hypothetical prote
15	27	100.0	498	2	T34592	DEAD-box RNA helic
16	27	100.0	500	2	B81060	N utilization subs
17	27	100.0	505	2	H81816	C2H2 zinc finger p
18	27	100.0	525	2	S34610	nitrogenase vanadi
19	27	100.0	529	2	S44649	f42h10.7 protein -
20	27	100.0	534	2	T24141	hypothetical prote
21	27	100.0	601	2	T35962	probable oxidoredu
22	27	100.0	730	2	A53064	folded gastrulatio
23	27	100.0	781	2	T50055	C2H2 zinc finger p
24	27	100.0	794	2	T09643	glutamine-tRNA lig
25	27	100.0	1278	2	T27925	hypothetical prote
26	26	96.3	71	2	B64479	hypothetical prote
27	26	96.3	110	2	C71657	hypothetical prote
28	26	96.3	126	2	T24181	hypothetical prote
29	26	96.3	134	1	WXBPT5	gene A2 protein -

ALIGNMENTS

RESULT 1

G82795

hypothetical protein XF0516 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: G82795

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide S

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: G82795

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-106 <SIM>

A:Cross-references: GB:AB003900; GB:AE003849; NID:G9105366; PIDN:AAF83326.1; GSPDB:

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.;

30	26	96.3	138	2	B46348	gene A2-A3 protein
31	26	96.3	228	2	H70181	ribose 5-phosphate
32	26	96.3	296	2	T34460	hypothetical prote
33	26	96.3	311	2	T40886	hypothetical prote
34	26	96.3	314	2	A38888	COI intron 8 prote
35	26	96.3	317	2	AH0351	hypothetical prote
36	26	96.3	349	2	AD3650	trans-1,2-dihydrob
37	26	96.3	365	2	H84001	aminomethyltransfe
38	26	96.3	379	2	AB1637	iron-sulfur cofact
39	26	96.3	431	2	B1927	probable oxidoredu
40	26	96.3	471	2	S05392	hypothetical prote
41	26	96.3	494	2	B98870	hypothetical prote
42	26	96.3	517	2	A13201	hypothetical prote
43	26	96.3	583	1	A41129	radixin - mouse
44	26	96.3	591	2	E71651	hypothetical prote
45	26	96.3	612	2	E97790	aminopeptidase hom
46	26	96.3	633	2	S19352	amino acid transpo
47	26	96.3	641	2	F97573	hypothetical prote
48	26	96.3	641	2	AG2794	hypothetical prote
49	26	96.3	659	2	S65976	yyr protein - Bac
50	26	96.3	664	2	AB1430	hypothetical prote
51	26	96.3	720	2	T51007	hypothetical prote
52	26	96.3	947	2	H84866	hypothetical prote
53	26	96.3	1098	2	JQ2209	helicase homolog 9
54	26	96.3	1133	2	JT0665	helicase II-like p
55	26	96.3	2145	2	S61041	glutamate synthase
56	24	88.9	82	2	T10934	cellulase (EC 3.2.
57	24	88.9	97	2	S72988	hypothetical prote
58	24	88.9	109	2	G87206	hypothetical prote
59	24	88.9	110	2	B83481	hypothetical prote
60	24	88.9	126	2	T42321	hypothetical prote
61	24	88.9	147	2	F70559	hypothetical prote
62	24	88.9	152	2	D81659	RabW, probable TCO
63	24	88.9	160	2	A71723	ribosomal protein
64	24	88.9	183	2	C75386	ribosome recycling
65	24	88.9	184	1	QCQVLG	BLI protein - toma
66	24	88.9	207	1	I64168	probable translati
67	24	88.9	215	2	F82062	adenylisulfate ki
68	24	88.9	225	1	MMIH68	E1 membrane glycop
69	24	88.9	231	2	E98256	hypothetical prote
70	24	88.9	231	2	A83028	hypothetical prote
71	24	88.9	232	2	AB3264	succinoglycan bios
72	24	88.9	236	2	C75181	hypothetical prote
73	24	88.9	269	2	B71365	probable phosphome
74	24	88.9	290	2	S44091	oxoglutarate/malat
75	24	88.9	290	2	G72858	AcOrf-70 protein -

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak
M.;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0516

Query Match 100.0%; Score 27; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 7 KHAVD 11

RESULT 2

T02603
hypothetical protein At2g14730 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F26C24.13
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
C;Accession: T02603; G84520
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
A;Reference number: Z14680
A;Accession: T02603
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-117 <ROU>
A;Cross-references: EMBL:AC004705; NID:G3252804; PID:G3252814
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84520
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-117 <STO>
A;Cross-references: GB:AE002093; NID:G3252814; PIDN:AAC24184.1; GSPDB:GN00139
C;Genetics:
A;Gene: F26C24.13; At2g14730
A;Map position: 2
A;Introns: 18/1; 80/1

Query Match 100.0%; Score 27; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 17 KHAVD 21

RESULT 3

F64214
histidine triad protein MG132 - Mycoplasma genitalium

C;Species: Mycoplasma genitalium
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999
C;Accession: F64214
R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhrman, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
C.A.; Venter, J.C
Science 270, 397-403, 1995

A;Title: The minimal gene complement of Mycoplasma genitalium.
A;Reference number: A64200; MUID:96026346; PMID:7569993
A;Accession: F64214

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-141 <TIGR>

A;Cross-references: GB:U39692; GB:L43967; NID:gl045812; PID:gl045813; TIGR:MG132
A;Experimental source: strain G-37

C;Genetics:
A;Genetic code: SGC3
C;Superfamily: protein kinase C inhibitor; histidine triad homology
F;9-105/Domain: histidine triad homology <HIT>
F;102-106/Region: histidine triad motif

Query Match 100.0%; Score 27; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 50 KHAVD 54

RESULT 4

T05656

hypothetical protein F22113.50 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 04-Mar-2000
C;Accession: T05656
R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15420
A;Accession: T05656
A;Molecule type: DNA
A;Residues: 1-159 <BEV>
A;Cross-references: EMBL:AL035539
A;Experimental source: cultivar Columbia; BAC clone F22113
C;Genetics:

A;Map position: 4

A;Introns: 117/1

A;Note: F22113.50

C;Superfamily: Arabidopsis thaliana hypothetical protein F22113.50

Query Match 100.0%; Score 27; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 128 KHAVD 132

RESULT 5

E89972

hypothetical protein [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: E89972

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: E89972

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-199 <KUR>

A;Cross-references: GB:BA000018; PID:gl3701647; PIDN:BA842940.1; GSPDB:GN00149
A;Experimental source: strain N315

C;Genetics:

A;Gene: SA1671

Query Match 100.0%; Score 27; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 55 KHAVD 59

RESULT 6

C84888
hypothetical protein At2g45250 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C:Accession: C84888
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
euss, D.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84888
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <STO>
A:Cross-references: GB:AE002093; NID:g2583136; PIDN:AAB82645.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g45250
A:Map position: 2
C:Superfamily: Arabidopsis thaliana hypothetical protein F22113.50

Query Match 100.0%; Score 27; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 169 KHAVD 173

RESULT 7

D64306
hypothetical protein MJ0052 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Oct-1999
C:Accession: D64306
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: D64306
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-222 <BUL>
A:Cross-references: GB:U67463; GB:L77117; NID:g1590846; PIDN:AAB98033.1; PID:g1498813; T
C:Genetics:
A:Map position: REV53374-52706

Query Match 100.0%; Score 27; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 217 KHAVD 221

RESULT 8

H71032
hypothetical protein PHI552 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: H71032
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; O.
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; O.
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: H71032
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <KAW>
A:Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30664.1; PID:g3257981
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Gen
C:Genetics:
A:Gene: PHI552

Query Match 100.0%; Score 27; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 110 KHAVD 114

RESULT 9

H84422
probable ribose 5-phosphate isomerase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: H84422
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallo
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84422
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <STO>
A:Cross-references: GB:AE002093; NID:g4262236; PIDN:AAD14529.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g01290
A:Map position: 2
C:Superfamily: Haemophilus influenzae ribose-5-phosphate isomerase

Query Match 100.0%; Score 27; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 63 KHAVD 67

RESULT 10

F87464
conserved hypothetical protein CC1738 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: F87464
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.;
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: F87464
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-382 <STO>
A;Cross-references: GB:AE005673; NID:gl3423156; PIDN:AAK23714.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC1738
C;Superfamily: Helicobacter pylori conserved hypothetical protein HP1020

Query Match 100.0%; Score 27; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 302 KHAVD 306

RESULT 11

Ai3490
membrane-bound lytic murein transglycosylase b (EC 3.2.1.1-) [imported] - Brucella melite
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AI3490
R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AI3490
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL53092.1; PID:gl7983956; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BM11911
A;Map position: 1
C;Keywords: glycosidase; hydrolase

Query Match 100.0%; Score 27; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 223 KHAVD 227

RESULT 12

G81179
oxidoreductase, probable NMB0614 [imported] - Neisseria meningitidis (strain MC58 serog
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81179
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81179
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-431 <TET>
A;Cross-references: GB:AE002416; GB:AE002098; NID:G7225832; PIDN:AAF41041.1; PID:G722584
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0614
C;Superfamily: hypothetical protein HT0499

Query Match 100.0%; Score 27; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 118 KHAVD 122

RESULT 13

S16867
gene H5 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C;Accession: S16867
R;Kato, K.
submitted to the EMBL Data Library, August 1991
A;Description: Sequence analysis of twenty mouse brain cDNA clones selected by spec
A;Reference number: S16416
A;Accession: S16867
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-478 <KAT>
A;Cross-references: EMBL:X61452; NID:G51202; PIDN:CAA43692.1; PID:G51203

Query Match 100.0%; Score 27; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 190 KHAVD 194

RESULT 14

A95278
hypothetical protein SMA0244 [imported] - Sinorhizobium meliloti (strain 1021) maga
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: A95278
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.;
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yen,
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium me
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: A95278
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-481 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK64787.1; PID:gl4523196; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSYMA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hu
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher,
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Le
heault, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Ye
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMA0244
A;Genome: plasmid
C;Superfamily: glycolate oxidase chain glcd

Query Match 100.0%; Score 27; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||||
DB 464 KHAVD 468

RESULT 15

T34592
DEAD-box RNA helicase - Streptomyces coelicolor

```

C;Cross-references: EMBL:U08403; NAD:G2898/4; FID:G2898/7
C;Genetics:
A;Introns: 67/3; 154/3; 212/3; 411/2
Query Match 100.0%; Score 27; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVD 5
|||
Db 229 KHAVD 233
RESULT 20
T24141
hypoethical protein R1058.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text change 15-Oct-1999

```

C:Accession: T24141
R:Ainecough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19844
A:Accession: T24141
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-534 <WIL>
A:Cross-references: EMBL:Z81576; PIDN:CAB04641.1; GSPDB:GN00023; CESP:R10E8.4
A:Experimental source: Clone R10E8
C:Genetics:
A:Gene: CESP:R10E8.4
A:Map position: 5
A:Introns: 200/3; 277/3; 311/2; 416/1; 461/3

Query Match 100.0%; Score 27; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||
DB 419 KHAVD 423

RESULT 21

T35962
probable oxidoreductase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: T35962
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z21551
A:Accession: T35962
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-601 <SER>
A:Cross-references: EMBL:AL035161; PIDN:CAA22723.1; GSPDB:GN00070; SCOEDB:SC9C7.11C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC9C7.11C
A:Superfamily: 3-hydroxyacyl-CoA dehydrogenase homology
F:13-293/Domain: 3-hydroxyacyl-CoA dehydrogenase homology <HCD1>
F:319-599/Domain: 3-hydroxyacyl-CoA dehydrogenase homology <HCD2>

Query Match 100.0%; Score 27; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||
DB 492 KHAVD 496

RESULT 22

A53064

folded gastrulation (fog) precursor - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: A53064
R:Costa, M.; Wilson, E.T.; Wieschaus, E.
Cell 76, 1075-1089, 1994
A:Title: A putative cell signal encoded by the folded gastrulation gene coordinates cell
A:Reference number: A53064; MUID:94185167; PMID:8137424
A:Accession: A53064
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-730 <COS>
A:Cross-references: GB:U03717; NID:G430722; PIDN:AAA18955.1; PID:G430723
C:Genetics:
A:Gene: fog
A:Cross-references: FlyBase:FBgn0000719

Query Match 100.0%; Score 27; DB 2; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||
DB 349 KHAVD 353

RESULT 23

T50055

C2H2 zinc finger protein [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50055
R:Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1999
A:Reference number: Z25031
A:Accession: T50055
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-781 <HUN>
A:Cross-references: EMBL:AL133521; PIDN:CAB63541.1; GSPDB:GN00066; SPDB:SPAC1039.0E
A:Experimental source: strain 972h(-); cosmid ci039
C:Genetics:
A:Gene: SPDB:SPAC1039.05c
A:Map position: 1
A:Introns: 41/1; 416/3; 457/2; 557/1; 768/1

Query Match 100.0%; Score 27; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||
DB 662 KHAVD 666

RESULT 24

T09643

glutamine-tRNA ligase (EC 6.1.1.18) - yellow lupine

C:Species: Lupinus luteus (yellow lupine)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T09643
R:Statecka, M.; Rozek, M.; Barciszewski, J.; Mirande, M.
Eur. J. Biochem. 256, 80-87, 1998
A:Title: Modular evolution of the Glx-tRNA synthetase family. Rooting of the evolution
A:Reference number: Z16793; MUID:98417431; PMID:9746349
A:Accession: T09643
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-794 <SIA>
A:Cross-references: EMBL:X91787; NID:G2995454; PIDN:CAA62301.1; PID:G2995455
A:Experimental source: subsp. ventus
C:Superfamily: human glutamine-tRNA ligase; glutamine-tRNA ligase homology
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F:271-547/Domain: glutamine-tRNA ligase homology <EGL>

Query Match 100.0%; Score 27; DB 2; Length 794;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
|||
DB 131 KHAVD 135

RESULT 25

T27925

hypothetical protein ZKS93.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27925

R;McMurray, A.
 submitted to the EMBL Data Library, February 1996
 A;Reference number: Z20440
 A;Accession: T27925
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1278 <WIL>
 A;Cross-references: EMBL:Z69385; PIDN:CAA93427.1; GSPDB:GN000022; CESP:ZK593.5
 A;Experimental source: clone ZK593
 C;Genetics:
 A;Gene: CESP:ZK593.5
 A;Map position: 4
 A;Introns: 11/1; 105/3; 139/3; 163/1; 232/3; 479/3; 505/2; 684/3; 828/3; 909/1; 1048/2;

Query Match 100.0%; Score 27; DB 2; Length 1278;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 |||||
 Db 558 KHAVD 562

Search completed: June 20, 2003, 20:58:58
 Job time : 16.7273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 / Search time 6.02273 Seconds
(without alignments)
34.433 Million cell updates/sec

Title: US-10-105-008-12

Perfect score: 27

Sequence: 1 KHAVD 5

Scoring table: BLOSUM62
Gapop 10.0 / Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	141	1 YHIT MYCGE	P47378 mycoplasma
2	27	100.0	222	1 Y052 METJA	Q0359 methanococ
3	27	100.0	236	1 YF52 PYRHO	O59220 pyrococcus
4	27	100.0	265	1 RPIA-ARATH	Q92438 arabidopsia
5	27	100.0	382	1 ISDF-CAUCR	Q94715 c ispd/ispi
6	27	100.0	478	1 SEB4 MOUSE	O43236 homo sapien
7	27	100.0	478	1 SEP4 MOUSE	P28661 mus musculu
8	27	100.0	529	1 YL27 CAEEL	P34420 caenorhabdi
9	27	100.0	730	1 FOG DROME	P40795 drosophila
10	27	100.0	794	1 SYO_LUEPLU	P52780 lupinus lut
11	26	96.3	71	1 YE35 METJA	Q58830 methanococ
12	26	96.3	110	1 Y532 RICPR	Q92418 rickettsia
13	26	96.3	134	1 VA2 BPT5	P23541 bacterioph
14	26	96.3	138	1 VA23 BPF2	P19348 bacterioph
15	26	96.3	228	1 RPIA-BORBU	O51601 borrelia bu
16	26	96.3	236	1 RPEC CAEEL	P91529 caenorhabdi
17	26	96.3	365	1 GCST-BACHD	Q9K334 bacillus ha
18	26	96.3	492	1 MURE STAAU	O86491 staphylococ
19	26	96.3	494	1 MURE STAAU	Q99V74 staphylococ
20	26	96.3	583	1 RADI MOUSE	P26043 mus musculu
21	26	96.3	633	1 AGP1 YEAST	P23776 saccharomyc
22	26	96.3	659	1 YETB-BACSU	P37484 bacillus su
23	26	96.3	2144	1 GLTI YEAST	Q12880 saccharomyc
24	24	88.9	153	1 RIBI RHIME	Q92400 rhizobium m
25	24	88.9	160	1 RS7 RICPR	P41081 rickettsia
26	24	88.9	183	1 RRF DEIRA	Q9RU82 deinococcus
27	24	88.9	199	1 SP25 DROME	Q9VY2 drosophila
28	24	88.9	207	1 YC10 HAEIN	P45103 haemophilus
29	24	88.9	215	1 CYSC-VIBCH	Q9K21 vibrio chol
30	24	88.9	225	1 VMEI ISV6	P05136 avian infec
31	24	88.9	236	1 Y418 PYRAB	Q9V121 pyrococcus
32	24	88.9	269	1 THID TREPA	O83153 treponema p
33	24	88.9	290	1 Y070 NPVAC	P41470 autographa

34	88.9	293	1	VBL1 ABMYW	P21946 abutilon mo
35	88.9	293	1	VBL1 BGMV	P06001 bean golden
36	88.9	293	1	VBL1 PYMVV	P27267 potato yell
37	88.9	293	1	VBL1 SLCV	P27267 squash leaf
38	88.9	293	1	VBL1 TGMV	P03566 tomato gold
39	88.9	293	1	VBL1 TMOV	Q05660 tomato mott
40	88.9	313	1	ACY2 HUMAN	P45381 homo sapien
41	88.9	320	1	YFHA-BORPE	Q93245 bordetella
42	88.9	342	1	QUEA-STRPV	P33449 streptococ
43	88.9	351	1	HOXN-ALCEU	P23516 alcaligenes
44	88.9	374	1	FLAI-BARBA	P35633 bartonella
45	88.9	381	1	HUPN-BRAJA	Q45247 bradyrhizob
46	88.9	416	1	PGK TRIRE	P14250 trichoderma
47	88.9	417	1	PGK TRIVI	P14250 trichoderma
48	88.9	462	1	NORM-BURVI	Q9F5N7 burkholderi
49	88.9	481	1	APRF-PSEAE	Q03027 pseudomonas
50	88.9	511	1	C772-SOLME	P37124 solanum mel
51	88.9	563	1	Y288-CHLTR	O84290 chlamydia t
52	88.9	576	1	CBPS-YEAST	P27614 saccharomyc
53	88.9	588	1	FRD2-SHEFR	Q924P0 shewanella
54	88.9	607	1	DB10-NICSV	P45942 nicotiana s
55	88.9	692	1	AMO-FICAN	P12807 pichia angu
56	88.9	692	1	UVRB-LACLA	Q9C106 lactococcus
57	88.9	696	1	EFG-RHILO	Q9N59 rhizobium l
58	88.9	752	1	HVSA-PROAC	Q59634 propionibac
59	88.9	907	1	ITH1-MOUSE	Q61702 mus musculu
60	88.9	960	1	UVRA-TREPA	O83527 treponema p
61	88.9	1069	1	ENTK-MOUSE	P97435 mus musculu
62	88.9	1072	1	HSEB-RAT	P23897 rattus norv
63	88.9	1256	1	ATL STAAU	P23897 rattus norv
64	88.9	1376	1	RPOB-RICTY	P77941 rickettsia
65	88.9	1517	1	YB22-SCHPO	Q10250 schizosacch
66	88.9	1820	1	CINA-ELEBL	P02719 electrophor
67	88.9	2164	1	CCAA-MOUSE	P97445 mus musculu
68	88.9	2212	1	CCAA-RAT	P54282 rattus norv
69	88.9	2327	1	CCAB-MOUSE	O55017 mus musculu
70	88.9	2336	1	CCAB-RAT	Q02294 rattus norv
71	88.9	2339	1	CCAB-HUMAN	Q00975 homo sapien
72	88.9	2339	1	CCAB-RABIT	Q05152 oryctolagus
73	88.9	2424	1	CCAA-RABIT	P27884 oryctolagus
74	88.9	2505	1	CCAA-HUMAN	O00555 homo sapien
75	85.2	40	1	Y557-HAEIN	P44015 haemophilus

ALIGNMENTS

RESULT 1				
YHIT MYCGE				
ID	YHIT MYCGE	STANDARD;	PRT;	141 AA.
AC	P47378;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical HIT-like protein MG132.			
GN	MG132.			
OS	Mycoplasma genitalium.			
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX	NCBI_Taxid=2037;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 33530 / G-37;			
RX	MEDLINE=96026346; PubMed=7569993;			
RA	Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,			
RA	Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,			
RA	Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,			
RA	Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,			
RA	Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Jacier T.S.,			
RA	Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;			
RT	"The minimal gene complement of Mycoplasma genitalium."			
RL	Science 270:397-403(1995).			
CC	-!- SIMILARITY: BELONGS TO THE HIT FAMILY.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U3693; AAC71349.1; -
DR HSSP; P49773; IKPF.

DR TIGR; MG132; -
DR InterPro; IPR001310; HIT.

DR Pfam; PF01230; HIT; 1.

DR PRINTS; PR00332; HISTRIAD.

DR PROSITE; PS00892; HIT; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 141 AA; 15587 MW; 0A0622386B036F2C CRC64;

Query Match 100.0%; Score 27; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
|
|
|
|
Db 50 KHAVD 54

RESULT 2

Y052 METJA

ID Y052 METJA

AC Q60359;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein M70052.

GN MJ0052.

OS Methanococcus jannaschii.

OC Archaea; Euryarchaeota; Methanococci; Methanococcales;

OC Methanocaldococcaceae; Methanocaldococcus.

OX NCBI_TaxID=2190;

RN [1]

RP SEQUENCE FROM N.A. 2661 / ATCC 43067;

RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

RX MEDLINE=96337999; PubMed=868087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,

RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,

RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,

RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,

RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,

RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;

RT "Complete genome sequence of the methanogenic archaeon, Methanococcus

jannaschii."

RL Science 273:1058-1073(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U67463; AAB98033.1; -

DR TIGR; MJ0052; -

DR InterPro; IPR001763; Rhodanese-like.

DR Pfam; PF00581; Rhodanese; 1.

DR SMART; SM00450; RHOD; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 222 AA; 25150 MW; 61991060852008A6 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 222;

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

Qy

Db

110 KHAVD 114

|
|
|
|

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fuji C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
Buell C.R., Ketchum K.A., Lee J.J., Rinning C.M., Koo H.L.,
Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
CC -1- CATALYTIC ACTIVITY: D-ribose 5-phosphate = D-ribulose 5-phosphate.
CC -1- PATHWAY: NONOXIDATIVE BRANCH OF THE PENTOSE PHOSPHATE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE RIBOSE 5-PHOSPHATE ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AC006200; AAD14529.1; -
DR InterPro: IPR004788; RpiA.
DR ProDom: PD005813; RpiA; 1.
DR TIGRFAMs: TIGR00021; rpiA; 1.
KW Hypothetical protein; Isomerase.
SQ SEQUENCE 265 AA; 28516 MW; 4A6587E6543691C1 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 265;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 63 KHAVD 67
|||||

RESULT 5
ISDF CAUCR STANDARD; PRT; 382 AA.
ID ISDF CAUCR STANDARD; PRT; 382 AA.
AC Q9A715;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE IspD/IspF bifunctional enzyme [Includes: 2-C-methyl-D-erythritol 4-
DE phosphate cyclidyltransferase (EC 2.7.7.60) (4-diphosphocyclidyl-2C-
DE methyl-D-erythritol synthase) (MEP cyclidyltransferase) (MCT); 2-C-
DE methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12)
DE (MECPs) (MECDP-synthase)].
GN ISPDF OR CC1738.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,

RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of *Caulobacter crescentus*.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: Bifunctional enzyme that catalyzes the formation of 4-
CC diphosphocyclidyl-2C-methyl-D-erythritol from CTP and 2C-methyl-D-
CC erythritol 4-phosphate (ispp), and converts 4-diphosphocyclidyl-2C-
CC methyl-D-erythritol 2-phosphate into 2C-methyl-D-erythritol 2,4-
CC cyclodiphosphate and CMP (ispp). Also converts 4-diphosphocyclidyl-
CC 2C-methyl-D-erythritol into 2C-methyl-D-erythritol 3,4-
CC cyclophosphate and CMP (ispp) (By similarity).
CC -1- CATALYTIC ACTIVITY: CTP + 2-C-methyl-D-erythritol 4-phosphate =
CC diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol.
CC -1- CATALYTIC ACTIVITY: 2-phospho-4-(cytidine 5'-diphospho)-2-C-
CC methyl-D-erythritol = 2-C-methyl-D-erythritol 2,4-cyclodiphosphate
CC + CMP.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; third step.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; fifth step.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ISPF FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ISPF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE005847; AAK23714.1; -
DR HSSP: Q46893; IINJ.
DR TIGR: CC1738; -
DR InterPro: IPR001228; ISPD synthase.
DR InterPro: IPR004423; Ispp.
DR InterPro: IPR003526; Ygbb.
DR Pfam: PF01128; ISPD; 1.
DR Pfam: PF02542; Ygbb; 1.
DR TIGRFAMs: TIGR00151; ispp; 1.
DR PROSITE: PS01295; ISPD; 1.
DR PROSITE: PS01350; ISPF; 1.
KW Transferase; Nucleotidyltransferase; Lyase; Isoprene biosynthesis;
KW Multifunctional enzyme; Complete proteome.
FT DOMAIN 1 222
FT 2-C-METHYL-D-ERYTHRITOL 2,4-
FT 4-DIPHOSPHOCYCLIDYL-2C-METHYL-D-ERYTHRITOL
FT SYNTHASE.
FT CYCLODIPHOSPHATE SYNTHASE.
SQ SEQUENCE 382 AA; 40048 MW; EZE0773856A317C CRC64;

Query Match 100.0%; Score 27; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
Db 302 KHAVD 306
|||||

RESULT 6
SEP4 HUMAN STANDARD; PRT; 478 AA.
ID SEP4 HUMAN STANDARD; PRT; 478 AA.
AC Q43236; Q9UM58;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Septin 4 (Peanut-like protein 2) {Brain protein HS} (Cell division
DE control-related protein 2) (HDCREL-2) (Bradeion beta) (CESB3 beta)
DE (Cerebral protein-7) (hucep-7).
GN SEPT4 OR PNUTL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN

RP SEQUENCE FROM N.A.
RA Zha D., Hu G.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99107819; PubMed=9889007;
RA Paavola P., Horelli-Kuitunen N., Palotie A., Peltonen L.;
RT "Characterization of a novel gene, PNUTL2, on human chromosome
17q22-q23 and its exclusion as the Meckel syndrome gene.";
RL Genomics 55:122-125(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC TISSUE=Brain;
RA Zieger B., Ware J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21402352; PubMed=11511094;
RA Tanaka M., Tanaka T., Kijima H., Itoh J., Matsuda T., Hori S.,
RY Yamamoto M.;
RT "Characterization of tissue- and cell-type-specific expression of a
novel human septin family gene, Bradeion.";
RL Biochem. Biophys. Res. Commun. 286:547-553(2001).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Brain;
RA Yoshimoto M., Yazaki M., Matsumoto K., Takayama K.;
RT "Molecular cloning of a new GTP binding protein from human brain.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN CYTOKINESIS (POTENTIAL).
CC -!- SUBUNIT: MAY ASSEMBLE INTO A MULTICOMPONENT STRUCTURE.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE SEPTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF035811; AAC88512.1; -;
CC EMBL; AF073312; AAC25673.1; -;
CC EMBL; U88829; AAD00653.1; -;
CC EMBL; U88870; AAD00657.1; -;
CC EMBL; AB008753; BAB70695.1; -;
CC EMBL; D89278; BAB69222.1; -;
CC EMBL; BC018056; AAH18056.1; -;
CC Genew; HGNC:9165; PNUTL2.
CC MIM; 603696; -;
CC InterPro; IPR000038; GTP_CDC; 1;
CC ProDom; PD002565; GTP_CDC; 1;
CC Cell division; GTP-binding; Coiled coil; Alternative splicing.
FT NP_BIND 151 158 GTP (POTENTIAL).
FT DOMAIN 447 478 COILED COIL (POTENTIAL).
FT VARSPLIC 1 20 MDRSLGQGNVPEPRTAG -> M (IN ISOFORM B).
SQ SEQUENCE 478 AA; 55098 MW; 2F08D3611EF6523D CRC64;

Query Match 100.0%; Score 27; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVD 5
|||||

Db 190 KHAVD 194
RESULT 7
SEP4_MOUSE STANDARD; PRT; 478 AA.
AC P28661;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SEPTIN 4 (Peanut-like protein 2) (Brain protein H5).
GN SEP4 OR PNUTL2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RA Kato K.;
RT "A collection of cDNA clones with specific expression patterns in
mouse brain.";
RL Eur. J. Neurosci. 2:704-711(1991).
CC -!- FUNCTION: INVOLVED IN CYTOKINESIS (POTENTIAL).
CC -!- SUBUNIT: MAY ASSEMBLE INTO A MULTICOMPONENT STRUCTURE.
CC -!- TISSUE SPECIFICITY: ABUNDANT IN AREAS OF HIGH CELL DENSITY,
PARTICULARLY IN THE STRIA TERMINALIS.
CC -!- SIMILARITY: BELONGS TO THE SEPTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X61452; CAA43692.1; -;
CC PIR; S16867; S16867.
CC MGD; MGI:1270156; Sept4.
CC InterPro; IPR000038; GTP_CDC; 1;
CC Pfam; PF00735; GTP_CDC; 1;
CC ProDom; PD002565; GTP_CDC; 1;
CC Cell division; GTP-binding; Coiled coil.
FT NP_BIND 151 158 GTP (POTENTIAL).
FT DOMAIN 447 478 COILED COIL (POTENTIAL).
SQ SEQUENCE 478 AA; 54936 MW; 36655DE54A6C4908 CRC64;
Query Match 100.0%; Score 27; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KHAVD 5
|||||
Db 190 KHAVD 194
RESULT 8
YL27_CAEEL STANDARD; PRT; 529 AA.
AC P34420;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein F42H10.7 in chromosome III.
GN F42H10.7
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copssey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laster N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Snaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohlman P.;
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L08403; AAA28021.1; -;
 DR WormPep; F42H10.7; CB00169.
 KW Hypothetical protein.
 SQ SEQUENCE 529 AA; 58284 MW; BBBP2C479BFE253 CRC64;

 QY 1 KHAVD 5 Query Match 100.0%; Score 27; DB 1; Length 529;
 DB 229 KHAVD 233 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 RESULT 9
 ID FOG_DROME STANDARD; PRT; 730 AA.
 AC P40795; Q9VVR71;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Folded gastrulation protein precursor.
 GN FOG OR G9559.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94185167; PubMed=8137424;
 RA Costa M., Wilson E.T., Wieschaus E.;
 RT "A putative cell signal encoded by the folded gastrulation gene
 RT coordinates cell shape changes during Drosophila gastrulation.";
 RL Cell 76:1075-1089(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Buseam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foster C., Gabriellian A.S., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svarksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -----
 CC -!- FUNCTION: COORDINATES CELL SHAPE CHANGES DURING FORMATION OF THE
 CC VENTRAL FURROW AND INVAGINATION OF THE POSTERIOR MIDGUT
 CC PRIMORDIUM, BY INDUCING APICAL CONSTRICTION OF CELLS IN SPATIALLY
 CC AND TEMPORALLY DEFINED MANNERS. FOG COULD FUNCTION AS A SECRETED
 CC SIGNAL TO INITIATE APICAL CONSTRICTION BY ACTING AS A LIGAND FOR
 CC AN UNIDENTIFIED G PROTEIN-COUPLED RECEPTOR, WHICH IN TURN
 CC ACTIVATES THE G PROTEIN ALPHA SUBUNIT ENCODED BY CONCERTINA, IN
 CC NEIGHBORING CELLS. SUCH AN INTRACELLULAR PATHWAY WOULD ULTIMATELY
 CC INDUCE CONTRACTION OF THE APICAL ACTIN-MYOSIN NETWORK. IN THE
 CC VENTRAL FURROW, FOG APPEARS TO ENSURE THAT ALL THE CELLS INITIATE
 CC CONSTRICTION WITHIN SEVERAL MINUTES OF EACH OTHER. IN THE
 CC POSTERIOR MIDGUT INVAGINATION, FOG APPEARS TO DIRECT THE ORDERED
 CC PROGRESSION OF CONSTRICTION INITIATIONS OUT FROM A CENTRAL REGION
 CC AND ALSO TO DELIMIT THE PERIPHERAL EXTENT OF THIS SPREADING.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE INVAGINATION PRIMORDIA IN A
 CC PATTERN THAT PRECISELY PRECEDES THE PATTERN OF CONSTRICTIONS.
 CC -!- DEVELOPMENTAL STAGE: MATERNAL TRANSCRIPTS ARE DEPOSITED INTO THE
 CC EGG AND UNIFORMLY DISTRIBUTED THROUGHOUT THE CORTEX OF CLEAVAGE
 CC STAGE AND SYNCYTIAL BLASTODERM EMBRYOS. ZYGOTIC TRANSCRIPTION IS
 CC FIRST FOUND IN THE VENTRAL FURROW PRIMORDIUM DURING THE BEGINNING
 CC OF CELLULARIZATION, ABOUT 30 MIN BEFORE THE START OF CONSTRICTIONS
 CC ALSO EXPRESSED ABOUT 30 MIN BEFORE THE START OF CONSTRICTIONS IN
 CC THE POSTERIOR MIDGUT PRIMORDIUM. THE VENTRAL-MOST CELLS ARE LAST
 CC TO EXPRESS FOG.
 CC -!- INDUCTION: CONTROLLED BY ZYGOTIC PATTERNING GENES.
 CC -!- PTM: THIS PROTEIN MAY BE HIGHLY O-GLYCOSYLATED AT ITS SER/THR-RICH
 CC C-TERMINAL. THIS COULD ANCHOR THE PROTEIN IN THE EXTRACELLULAR
 CC MATRIX UPON SECRETION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U03717; AAA18955.1; -;
 DR EMBL; AE003573; AAF50935.1; -;
 DR FlyBase; FBgn000719; fog.
 DR Developmental protein; Signal; Gastrulation; Extracellular matrix;
 KW Glycoprotein. 1 21 POTENTIAL.
 FT SIGNAL 22 730 FOLDED GASTRULATION PROTEIN.
 FT CHAIN 57 167 G PROTEIN-COUPLED RECEPTOR LIGAND

```
FT DOMAIN 579 586 (POTENTIAL)
FT CARBOHYD 51 51 POLY-THR.
FT CARBOHYD 193 193 N-LINKED (GLNAC. . .) (POTENTIAL)
FT CARBOHYD 252 252 N-LINKED (GLNAC. . .) (POTENTIAL)
FT CARBOHYD 289 289 N-LINKED (GLNAC. . .) (POTENTIAL)
FT CARBOHYD 459 459 N-LINKED (GLNAC. . .) (POTENTIAL)
FT CARBOHYD 590 590 N-LINKED (GLNAC. . .) (POTENTIAL)
FT CARBOHYD 639 639 N-LINKED (GLNAC. . .) (POTENTIAL)
FT CARBOHYD 678 678 N-LINKED (GLNAC. . .) (POTENTIAL)
SQ SEQUENCE 730 AA; 78250 MW; D00D426139AB987C CRC64;

Query Match 100.0%; Score 27; DB 1; Length 730;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 349 KHAVD 353

RESULT 10
SYQ_LUPLU STANDARD; PRT; 794 AA.
AC P53780;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glutaminyl-tRNA synthetase (EC 6.1.1.18) (Glutamine--tRNA ligase)
DS (GlnRS).
OS Lupinus luteus (Yellow lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
OX NCBI_TaxID=3873;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ventus;
RA Stacecka M., Rozek M., Barciszewski J.;
RT "Isolation and characterization of a cDNA clone encoding a plant gene
of aminoacyl-tRNA synthetase."
RL (In) Plant Gene Register PGR95-103.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamine + tRNA(Gln) = AMP +
diphosphate + L-glutaminyl-tRNA(Gln).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X91787; CAA62901.1;
CC HSPF; P00962; 1GTR.
DR InterPro; IPR004514; GlnS.
DR InterPro; IPR000924; Gln_rRNA-synt_1c.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00749; tRNA-synt_1c; 1.
DR TIGRFAMs; TIGR00440; GlnS. 1.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 277 287 "HIGH" REGION.
FT SITE 505 509 "KMSKS" REGION.
FT BINDING 508 508 ATP (BY SIMILARITY).
SQ SEQUENCE 794 AA; 90482 MW; 4C7D02C0795268E1 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 794;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
```

```
DB 131 KHAVD 135

RESULT 11
YE35_METJA STANDARD; PRT; 71 AA.
AC Q58830;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1435.
GN MJ1435.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.L.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073 (1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U67584; AAB99450.1;
CC TIGR; MJ1435;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 71 AA; 8295 MW; AC1B6AF9651500EB CRC64;

Query Match 96.3%; Score 26; DB 1; Length 71;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 63 KHAID 67

RESULT 12
Y532_RICPR STANDARD; PRT; 110 AA.
AC Q9ZDI8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein RP532.
GN RP532.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia;
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
```



```
RA Sicheritz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria"; 140(1998).
RL Nature 396:133-140(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ235272; CAA14981.1; -.
FW Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 88 108 POTENTIAL.
SQ SEQUENCE 110 AA; 13002 MW; D175C78E35FB04DA CRC64;

Query Match 96.3%; Score 26; DB 1; Length 110;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 71 KHAID 75

RESULT 13
VA2 VA2 BPT5 STANDARD; PRT; 134 AA.
AC P23541;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE A2 protein.
OS Bacteriophage T5.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC T5-like viruses.
OX NCBI_TaxID=10726;
RN [1]
RP SEQUENCE.
RX MEDLINE=91282771; PubMed=2059212;
RA Snyder C.E. Jr.;
RT Amino acid sequence of the bacteriophage T5 gene A2 protein.;
RL Biochem. Biophys. Res. Commun. 177:1240-1246(1991).
RN [2]
RP SEQUENCE OF 1-28.
RX MEDLINE=82256600; PubMed=7049170;
RA Fox J.W., Barish A., Snyder C.E. Jr., Benzinger R.;
RT "Amino terminal sequence of the bacteriophage T5-coded gene A2
RT protein.";
RL Biochem. Biophys. Res. Commun. 106:265-269(1982).
CC -|- FUNCTION: MAY COMPETE WITH THE HOST OMPF PROTEIN FOR BINDING TO
CC LIPOPOLYSACCHARIDES.
CC -|- MISCELLANEOUS: IT IS ONE OF THE TWO PROTEINS ENCODED BY FIRST-STEP
CC -|- TRANSFER DNA.
CC -|- SIMILARITY: STRONG, TO THE PRODUCT OF GENE A2-A3 OF BACTERIOPHAGE
CC BF23.
DR PIR; JQ1034; WXBPT5.
KW DNA-binding.
SQ SEQUENCE 134 AA; 14200 MW; AD2BB0EC050F45F3 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 134;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 90 KHAID 94

RESULT 14
VA23 BP8F2 STANDARD; PRT; 138 AA.
AC P19348;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE A2-A3 protein.
GN A2-A3.
OS Bacteriophage BF23.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC T5-like viruses.
OX NCBI_TaxID=10707;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wild type;
RX MEDLINE=90320148; PubMed=2196743;
RA Wiest J.S., McCorquodale D.J.;
RT "Characterization of preearly genes in the terminal repetition of
RT bacteriophage BF23 DNA by nucleotide sequencing and restriction
RT mapping.";
RL Virology 177:745-752(1990).
RN [2]
RP REVISIONS.
RX MEDLINE=91112013; PubMed=1989394;
RA Wiest J.S., McCorquodale D.J.;
RT "Characterization of preearly genes in the terminal repetition of
RT bacteriophage BF23 DNA by nucleotide sequencing and restriction
RT mapping.";
RL Virology 180:861-861(1991).
CC -|- FUNCTION: PROTEINS A1 AND A2-A3 ARE NECESSARY FOR COMPLETION
CC OF PHAGE DNA TRANSFER. A2-A3 BINDS TO HOST RNA POLYMERASE AND
CC IS ALSO INVOLVED IN THE ABORTIVE MECHANISM THAT OCCURS AFTER
CC A HOST CELL HARBORING A COLLB PLASMID IS INFECTED WITH BP23+.
CC -|- SIMILARITY: STRONG, TO THE PRODUCT OF GENE A2 BACTERIOPHAGE T5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37095; AAA32179.1; -.
SQ SEQUENCE 138 AA; 14672 MW; 4D5ECC911778E7A9 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 138;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 94 KHAID 98

RESULT 15
RPIA BORBU STANDARD; PRT; 228 AA.
AC O51601;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribose 5-phosphate isomerase A (EC 5.3.1.6) (Phosphoriboisomerase A).
GN RPIA OR B80657.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
```

RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Usterback T., Watthey L., McDonald L., Artach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi";
RL Nature 390:580-586 (1997)
CC -1- CATALYTIC ACTIVITY: D-ribose 5-phosphate = D-ribose 5-phosphate.
CC -1- PATHWAY: NONOXIDATIVE BRANCH OF THE PENTOSE PHOSPHATE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE RIBOSE 5-PHOSPHATE ISOMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; AE001167; AAC67008.1; -
DR TIGR; BB0657; -
DR InterPro; IPR004788; RpiA.
DR ProDom; PD005813; RpiA; 1.
DR TIGRFAMs; TIGR00021; rpiA; 1.
DR Isomerase; Complete proteome.
KW SEQUENCE 228 AA; 25885 MW; 4051FC370FF7C76F CRC64;
SQ
Query Match 96.3%; Score 26; DB 1; Length 228;
Best Local Similarity 80.0%; Pred. No. 38;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAUD 5
Db 10 KHAID 14
RESULT 16
RPC6 CAEL STANDARD; PRT; 296 AA.
AC P91529;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable DNA-directed RNA polymerases III 39 kDa polypeptide (EC
DE 2.7.7.6) (RNA polymerase III C39 subunit).
GN W09C3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
CC
CC
RA Wansley P., Krane J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC RPC34/RPC39 RNA POLYMERASE
CC SUBUNIT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; U88178; AAC24412.1; -
DR WormPep; W09C3.4; CE14750.
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Usterback T., Watthey L., McDonald L., Artach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi";
RL Nature 390:580-586 (1997)
CC -1- CATALYTIC ACTIVITY: D-ribose 5-phosphate = D-ribose 5-phosphate.
CC -1- PATHWAY: NONOXIDATIVE BRANCH OF THE PENTOSE PHOSPHATE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE RIBOSE 5-PHOSPHATE ISOMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; AE001167; AAC67008.1; -
DR TIGR; BB0657; -
DR InterPro; IPR004788; RpiA.
DR ProDom; PD005813; RpiA; 1.
DR TIGRFAMs; TIGR00021; rpiA; 1.
DR Isomerase; Complete proteome.
KW SEQUENCE 228 AA; 25885 MW; 4051FC370FF7C76F CRC64;
SQ
Query Match 96.3%; Score 26; DB 1; Length 228;
Best Local Similarity 80.0%; Pred. No. 38;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAUD 5
Db 10 KHAID 14
RESULT 16
RPC6 CAEL STANDARD; PRT; 296 AA.
AC P91529;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable DNA-directed RNA polymerases III 39 kDa polypeptide (EC
DE 2.7.7.6) (RNA polymerase III C39 subunit).
GN W09C3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
CC
CC
RA Wansley P., Krane J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC RPC34/RPC39 RNA POLYMERASE
CC SUBUNIT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; U88178; AAC24412.1; -
DR WormPep; W09C3.4; CE14750.

KW Transferase; DNA-directed RNA polymerase; Transcription;
KW Nuclear protein.
SQ SEQUENCE 296 AA; 32754 MW; 18D2520E340770F8 CRC64;
Query Match 96.3%; Score 26; DB 1; Length 296;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAUD 5
Db 185 KHAID 189
RESULT 17
GCST_BACHD STANDARD; PRT; 365 AA.
ID GCST_BACHD
AC G9K934;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable aminomethyltransferase (EC 2.1.2.10) (Glycine cleavage system
DE T protein).
GN GCVT OR BH2816.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RA MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331 (2000).
CC -1- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC glycine (By similarity).
CC -1- CATALYTIC ACTIVITY: (6S)-tetrahydrofolate + S-
CC aminomethyldihydrofoloylprotein = (6R)-5,10-
CC methylenetetrahydrofolate + NH(3) + dihydrofoloylprotein.
CC -1- SUBUNIT: THE GLYCINE CLEAVAGE SYSTEM IS COMPOSED OF FOUR PROTEINS:
CC P, T, L, AND H (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GCVT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; AF001516; BAB06535.1; -
DR InterPro; IPR002536; GCV.T.
DR Pfam; PF01571; GCV.T; 1.
DR TIGRFAMs; TIGR00528; gcvT; 1.
KW Transferase; Aminotransferase; Complete proteome.
SQ SEQUENCE 365 AA; 40309 MW; 867214D42B3691C5 CRC64;
Query Match 96.3%; Score 26; DB 1; Length 365;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAUD 5
Db 128 KHAID 132
RESULT 18
MURE STAAU STANDARD; PRT; 492 AA.
ID MURE STAAU
AC O86491;

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE UDP-N-acetylmuramoyl-L-glutamate-2,6-diaminopimelate ligase
DE (EC 6.3.2.13) (UDP-N-acetylmuramoyl-tripeptide synthetase) (Meso-
GN diaminopimelate-adding enzyme) (UDP-MurNac-tripeptide synthetase).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=98313013; PubMed=9650993;
RA Ludovice A.M., Wu S., de Lencastre H.;
RT "Molecular cloning and DNA sequencing of the Staphylococcus aureus
RT UDP-N-acetylmuramoyl tripeptide synthetase (murE) gene, essential for
RT the optimal expression of methicillin resistance."
RL Microb. Drug Resist. 4:85-90(1998).
CC -|- FUNCTION: CELL WALL FORMATION. DIAMINOPIMELIC ACID ADDING ENZYME
CC (BY SIMILARITY).
CC -|- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanyl-D-
CC glutamate + meso-2,6-diaminoheptanedioate = ADP + phosphate +
CC UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2,6-
CC diaminopimelate.
CC -|- PATHWAY: Peptidoglycan biosynthesis.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- SIMILARITY: BELONGS TO THE MURCDEF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y14370; CAA74740.1; -
DR InterPro: IPR000713; Mur_ligase.
DR InterPro: IPR004101; Mur_ligase_C.
DR Pfam: PF01225; Mur_ligase; 1.
DR Pfam: PF02875; Mur_ligase_C; 1.
DR TIGRFAMs: TIGR01085; murE; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
ATP-binding.
FT NP_BIND 109 115 ATP (POTENTIAL).
SQ SEQUENCE 492 AA; 54099 MW; 68873421B481FF1B CRC64;

Query Match 96.3%; Score 26; DB 1; Length 492;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 439 KHAID 443

RESULT 19
MURE STAA
ID MURE STAA
AC Q99V74; PRT; 494 AA.
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE UDP-N-acetylmuramoyl-L-glutamate-2,6-diaminopimelate ligase
DE (EC 6.3.2.13) (UDP-N-acetylmuramoyl-tripeptide synthetase) (Meso-
DE diaminopimelate-adding enzyme) (UDP-MurNac-tripeptide synthetase).
GN MURE OR SAV1018 OR SA0876.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratsugu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
CC -|- FUNCTION: Cell wall formation. Diaminopimelic acid adding enzyme
CC (BY SIMILARITY).
CC -|- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanyl-D-
CC glutamate + meso-2,6-diaminoheptanedioate = ADP + phosphate +
CC UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2,6-
CC diaminopimelate.
CC -|- PATHWAY: Peptidoglycan biosynthesis.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- SIMILARITY: BELONGS TO THE MURCDEF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AP003361; BAB57180.1; -
DR EMBL: AP003132; BAB42117.1; -
DR InterPro: IPR000713; Mur_ligase.
DR InterPro: IPR004101; Mur_ligase_C.
DR Pfam: PF01225; Mur_ligase; 1.
DR Pfam: PF02875; Mur_ligase_C; 1.
DR TIGRFAMs: TIGR01085; murE; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
ATP-binding; Complete proteome.
FT NP_BIND 110 116 ATP (POTENTIAL).
SQ SEQUENCE 494 AA; 54232 MW; B5BD224B23DA5A13 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 494;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 440 KHAID 444

RESULT 20
RADI MOUSE
ID RADI MOUSE STANDARD; PRT; 583 AA.
AC P26043;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Radixin.
GN RDX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-53 AND 263-277.
RC TISSUE=Liver.
RX MEDLINE=92064635; PubMed=1955455;
RA Funayama N., Nagafuchi A., Sato N., Tsukita S., Tsukita S.;
RT "Radixin is a novel member of the band 4.1 family."
RL J. Cell Biol. 115:1039-1048(1991).
CC -|- FUNCTION: PROBABLY PLAYS A CRUCIAL ROLE IN THE BINDING OF THE

BARBED END OF ACTIN FILAMENTS TO THE PLASMA MEMBRANE.
-1- SUBCELLULAR LOCATION: HIGHLY CONCENTRATED IN THE UNDERCOAT OF THE CELL-TO-CELL ADHERENS JUNCTION AND THE CLEAVAGE FURROW IN THE INTERPHASE AND MITOTIC PHASE, RESPECTIVELY.
-1- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X60672; CAA43087.1; -;
PIR; A41129; A41129.
MGD; MGI:97887; Rdx.
InterPro; IPR000299; Band 4.1.
InterPro; IPR000798; E2/rad/moesin.
Pfam; PF00373; Band_41; 1.
Pfam; PF00769; ERM; 1.
PRINTS; PR00935; BAND41.
SMART; SM00295; B41; 1.
PROSITE; PS00660; BAND_41_1; 1.
PROSITE; PS00661; BAND_41_2; 1.
PROSITE; PS00662; BAND_41_3; 1.
Structural protein; Cytoskeleton; Actin-binding; Capping protein.
DOMAIN 58 225 POLY-PRO.
DOMAIN 470 477
SEQUENCE 583 AA; 68451 MW; 3219A52ECDDB7EBE CRC64;

Query Match 96.3%; Score 26; DB 1; Length 583;
Best Local Similarity 80.0%; Pred. No. 1.e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 574 KHAID 578

RESULT 21
AGPI YEAST
ID AGPI YEAST STANDARD; PRT; 633 AA.
AC P25376;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Asparagine/glutamine permealase.
GN AGPI OR YCL025C OR YCL25C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Hollenberg C.P., Kleinhaus U., Luetzenkirchen K., Rad M.R., Xu G.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION.
RA Schreve J.L., Sin J., Garrett J.M.;
RL Unpublished observations (JUL-1997).
-1- FUNCTION: BROAD SUBSTRATE RANGE PERMEASE WHICH TRANSPORTS ASPARAGINE AND GLUTAMINE WITH INTERMEDIATE SPECIFICITY.
-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
-1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X59720; CAA42360.1; -;
DR PIR; S19352; S19352.
DR SGD; S0000530; AGPI.
DR InterPro; IPR002293; AA/rel_primease1.
DR InterPro; IPR004840; AAC_permealase.
DR InterPro; IPR004841; Permealase.
DR InterPro; IPR004762; Yeast_AA_perm.
DR Pfam; PF00324; aa_permealases; 1.
DR TIGRFAMs; TIGR00913; ZAO310; 1.
KW PROSITE; PS00218; AMINO-ACID PERMEASE 1; 1.
KW TRANSPORT; Amino-acid transport; Transmembrane.
FT TRANSMEM 125 141 POTENTIAL.
FT TRANSMEM 152 169 POTENTIAL.
FT TRANSMEM 191 214 POTENTIAL.
FT TRANSMEM 236 252 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
FT TRANSMEM 354 372 POTENTIAL.
FT TRANSMEM 407 425 POTENTIAL.
FT TRANSMEM 454 471 POTENTIAL.
FT TRANSMEM 478 499 POTENTIAL.
FT TRANSMEM 525 549 POTENTIAL.
FT TRANSMEM 565 581 POTENTIAL.
SQ SEQUENCE 633 AA; 69812 MW; 05D1F744AP939AE7 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 633;
Best Local Similarity 80.0%; Pred. No. 1.e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 306 KHAID 310

RESULT 22
YVBT BACSU
ID YVBT BACSU STANDARD; PRT; 659 AA.
AC P37484;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yvbt.
GN YVBT
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogawawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chromosome containing the replication origin."
RL DNA Res. 1:1-14(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst P., Ogawawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerston I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Hachez J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,"
RL Nature 390:249-256 (1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D26185; BAA05182.1; --
DR EMBL; Z59124; CAB16088.1; --
DR Subtilisin; BG10010; Yybt.
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR00160; GGDEF.
DR InterPro; IPR001667; Ppesterase.
DR Pfam; PF01368; DHH; 1.
DR Pfam; PF02272; DHHA1; 1.
DR SMART; SM00267; DUF1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 659 AA; 74325 MW; FD325C28A10EB14B CRC64;
Query Match 96.3%; Score 26; DB 1; Length 659;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAUD 5
DB 646 KHAID 650
|||:|
RESULT 23
GLT1_YEAST
ID GLT1_YEAST STANDARD; PRT; 2144 AA.
AC Q12680; Q12290;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Glutamate synthase [NADPH] precursor (EC 1.4.1.13) (NADPH-GOGAT).
GN GLT1 OR YD117C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CN36;
RX MEDLINE=97082505; PubMed=8923741;
RA Filetici P., Martegani M.P., Valenzuela L., Gonzalez A., Ballario P.;
RT "Sequence of the GLT1 gene from *Saccharomyces cerevisiae* reveals the
RT domain structure of yeast glutamate synthase.";
RL Yeast 12:1359-1366 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: 2 L-glutamate + NADP(+) = L-glutamine + 2-
CC oxoglutarate + NADPH
CC -1- COFACTOR: BINDS A 3FE-4S CLUSTER; FAD AND FMN.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SIMILARITY: TO OTHER GLUTAMATE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X89221; CAA61505.1; --
DR EMBL; Z67750; CAA91574.1; --
DR EMBL; Z74219; CAA98745.1; --
DR SGD; S0002330; GLT1.
DR InterPro; IPR002489; DUF14.
DR InterPro; IPR001327; FAD pyr redox.
DR InterPro; IPR002932; Glu_synthase.
DR Pfam; PF00070; Pyr_redox; 1.
DR Pfam; PF01493; DUF14; 1.
DR Pfam; PF01645; Glu synthase; 1.
DR ProDom; PD000139; FAD pyr redox; 1.
DR TIGRfam; TIGR01117; GOGAT sm gam; 1.
KW Oxidoreductase; Iron-sulfur; 3Fe-4S; Flavoprotein; FAD; FMN; NADP;
KW Glutamate biosynthesis; Zymogen.
FT PROPEP 1 53
FT CHAIN 54 2144 GLUTAMATE SYNTHASE [NADPH].
FT NP BIND 1131 1183 FMN (BY SIMILARITY).
FT METAL 1184 1184 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
FT METAL 1190 1190 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
FT METAL 1195 1195 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
FT CONFLICT 30 30 L -> D (IN REF. 2).
FT CONFLICT 166 172 TSRRFY -> NVPVDSTI (IN REF. 2).
FT CONFLICT 449 451 IPS -> FLV (IN REF. 2).
FT CONFLICT 1752 1752 L -> V (IN REF. 2).
SQ SEQUENCE 2144 AA; 238200 MW; 5AAG6A948EF955349 CRC64;
Query Match 96.3%; Score 26; DB 1; Length 2144;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAUD 5
DB 772 KHAID 776
|||:|
RESULT 24
RIBI_RHIME
ID RIBI_RHIME STANDARD; PRT; 153 AA.
AC Q92Q00;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase 1 (EC 2.5.1.9) (DMRL synthase
DE 1) (lumazine synthase 1) (Riboflavin synthase 1 beta chain).
GN RIB1 OR R01215 OR SMC01777.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Anpe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Gouzie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
 CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
 CC ribityl-amino-2,4(H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
 CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
 CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
 CC 2,4(H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-
 CC phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =
 CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
 CC -1- PATHWAY: Riboflavin biosynthesis; last step.
 CC -1- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AL591786; CAC45794.1; -;
 DR InterPro; IPR002180; DMRL_synthase.
 DR Pfam; PF00885; DMRL_synthase; 1.
 DR ProDom; PD003664; DMRL_synthase; 1.
 KW Riboflavin biosynthesis; Transferase; Complete proteome.
 SQ SEQUENCE 153 AA; 16193 MW; 8F50CA43DC559856 CRC64;

Query Match 88.9%; Score 24; DB 1; Length 153;
 Best Local Similarity 80.0%; Pred. No. 76;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
 |||:
 Db 29 KHALD 33

RESULT 25
 RS7_RICPR STANDARD; PRT; 160 AA.
 AC P41081;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 30S ribosomal protein S7.
 GN RPSG OR RP131.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsieae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RA Wood D.O.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O., Naeslund A.K.,
 RA Sichertitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria."
 RL Nature 396:133-146(1998).
 CC -1- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
 CC 16S RIBOSOMAL RNA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U02603; AAA18330.1; -;
 DR EMBL; AJ235270; CAA14600.1; -;
 DR HSPSP; P22744; IHUS.
 DR InterPro; IPR000235; Ribosomal_S7.
 DR Pfam; PF00177; Ribosomal_S7; 1.
 DR ProDom; PD000817; Ribosomal_S7; 1.
 DR TIGRFAMs; TIGR01029; rpsG bact; 1.
 DR PROSITE; PS00052; RIBOSOMAL_S7; 1.
 KW Ribosomal protein; rRNA-binding; Complete proteome.
 SQ SEQUENCE 160 AA; 18309 MW; 5C66702A37F41761 CRC64;

Query Match 88.9%; Score 24; DB 1; Length 160;
 Best Local Similarity 80.0%; Pred. No. 80;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
 |||:
 Db 53 KHSVD 57

Search completed: June 20, 2003, 20:52:55
 Job time : 10.0227 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:42:49 ; Search time 26.3636 Seconds
(without alignments)
39.078 Million cell updates/sec

Title: US-10-105-008-12

Perfect score: 27

Sequence: 1 KIAVD 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database :

SPTREMBL 21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	106	16 Q9PF20	Q9pf20 xylella fas
2	27	100.0	117	10 O80978	O80978 arabidopsis
3	27	100.0	159	10 Q9SVF7	Q9svf7 arabidopsis
4	27	100.0	179	16 Q9RKA9	Q9rka9 streptomyce
5	27	100.0	199	16 Q9RT24	Q9rt24 staphylococ
6	27	100.0	211	10 O22147	O22147 arabidopsis
7	27	100.0	266	5 Q9NEJ3	Q9nej3 caenorhabdi
8	27	100.0	274	4 Q9H315	Q9h315 homo sapien
9	27	100.0	293	5 O15988	O15988 patinopecte
10	27	100.0	296	5 Q95XP1	Q95xp1 caenorhabdi
11	27	100.0	314	5 O15987	O15987 patinopecte
12	27	100.0	356	4 Q9S651	Q9s651 homo sapien
13	27	100.0	415	16 Q98BM3	Q98bm3 rhizobium l
14	27	100.0	427	16 Q8YEG7	Q8yeg7 bruceella me
15	27	100.0	431	16 Q9K015	Q9k015 neisseria m
16	27	100.0	473	16 Q9LOF7	Q9lof7 streptomyce

ALIGNMENTS

RESULT 1
Q9PF20 PRELIMINARY; PRT; 106 AA.
ID Q9PF20
AC Q9PF20
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein Xf0516.
GN Xf0516.

17	27	100.0	481	16 Q93OR6	Q93or6 rhizobium m
18	27	100.0	498	16 O54116	O54116 streptomyce
19	27	100.0	500	16 Q9JYD3	Q9jyd3 neisseria m
20	27	100.0	505	16 Q9JTB6	Q9jtb6 neisseria m
21	27	100.0	525	2 Q46243	Q46243 clostridium
22	27	100.0	534	5 Q9XUR9	Q9xur9 caenorhabdi
23	27	100.0	535	2 Q9S0V3	Q9s0v3 nitrogen fi
24	27	100.0	601	16 Q9ZBJ9	Q9zbj9 streptomyce
25	27	100.0	656	2 Q54907	Q54907 streptococc
26	27	100.0	704	3 Q8X002	Q8x002 neurospora
27	27	100.0	781	3 Q9US36	Q9us36 schizosacch
28	27	100.0	1043	5 Q95TK6	Q95tk6 drosophila
29	27	100.0	1071	12 Q9YZA1	Q9yza1 gallid herp
30	27	100.0	1278	5 Q23542	Q23542 caenorhabdi
31	27	100.0	1655	5 Q9VB55	Q9vb55 drosophila
32	27	100.0	1888	5 Q9N9Z6	Q9n9z6 drosophila
33	26	96.3	107	12 Q9WL73	Q9wl73 dobrava vir
34	26	96.3	111	12 Q9YTC3	Q9ytc3 dobrava vir
35	26	96.3	126	5 Q21930	Q21930 caenorhabdi
36	26	96.3	161	12 Q8UZD3	Q8uzd3 cercopithec
37	26	96.3	167	12 Q9DJK8	Q9djks dobrava vir
38	26	96.3	167	12 Q9DJK7	Q9djks dobrava vir
39	26	96.3	167	12 Q9DJK6	Q9djks dobrava vir
40	26	96.3	179	12 Q9W7X6	Q9w7x6 dobrava vir
41	26	96.3	183	12 Q9W7T3	Q9w7t3 dobrava vir
42	26	96.3	184	12 Q72481	Q72481 dobrava vir
43	26	96.3	184	12 Q9W7X5	Q9w7x5 dobrava vir
44	26	96.3	185	12 Q89917	Q89917 dobrava vir
45	26	96.3	185	12 Q89918	Q89918 dobrava vir
46	26	96.3	185	12 Q89919	Q89919 dobrava vir
47	26	96.3	185	12 Q89920	Q89920 dobrava vir
48	26	96.3	185	12 Q89921	Q89921 dobrava vir
49	26	96.3	185	12 Q89922	Q89922 dobrava vir
50	26	96.3	185	12 Q9IEQ0	Q9ieq0 dobrava vir
51	26	96.3	186	12 Q9IEQ1	Q9ieq1 dobrava vir
52	26	96.3	186	12 Q9DJK5	Q9djks dobrava vir
53	26	96.3	236	12 Q91BC8	Q91bc8 spodoptera
54	26	96.3	296	5 Q9VDM3	Q9vdm3 drosophila
55	26	96.3	311	3 Q74848	Q74848 schizosacch
56	26	96.3	314	8 Q02683	Q02683 podospora a
57	26	96.3	317	5 Q9VTY2	Q9vty2 drosophila
58	26	96.3	317	5 Q8SXE8	Q8sxe8 drosophila
59	26	96.3	317	13 Q90W83	Q90w83 gallus gall
60	26	96.3	317	16 Q8ZCS8	Q8zcs8 versinia pe
61	26	96.3	319	13 Q9DDC7	Q9ddc7 lepidodacty
62	26	96.3	349	16 Q8YAX6	Q8yax6 bruceella me
63	26	96.3	370	16 Q98SH2	Q98sh2 rhizobium l
64	26	96.3	379	16 Q92BB5	Q92bb5 listeria in
65	26	96.3	425	5 Q9W3C3	Q9w3c3 drosophila
66	26	96.3	429	12 Q9YKX5	Q9ykk5 dobrava vir
67	26	96.3	429	12 Q9YQ49	Q9yq49 dobrava vir
68	26	96.3	429	12 Q9YPT7	Q9ypt7 dobrava vir
69	26	96.3	429	12 Q9YPT6	Q9ypt6 dobrava vir
70	26	96.3	429	12 Q9DJL0	Q9djlo dobrava vir
71	26	96.3	429	12 Q9DJL9	Q9djln dobrava vir
72	26	96.3	429	12 Q82136	Q82136 hantavirus
73	26	96.3	431	16 Q9JUV0	Q9jvjo neisseria m
74	26	96.3	475	2 Q93CT1	Q93ct1 angelia bo
75	26	96.3	517	16 Q8UJY1	Q8ujy1 agrobacteri

OS *Xylella fastidiosa*.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC *Xylella*.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Bonaccorsi A.A., Camargo L.E.A., Carraro D.M., Carver H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Pacinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lenos E.G.M., Lenos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Zetubal J.C.;
 RA "The genome sequence of the plant pathogen *Xylella fastidiosa*."
 RL Nature 406:151-159(2000).
 DR EMBL: AE003900; AAF83326.1; --
 KW Hypothetical protein; Complete proteome
 SQ SEQUENCE 106 AA; 11571 MW; 54D49E4CAF9D876 CRC64;
 Query Match 100.0%; Score 27; DB 16; Length 106;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVD 5
 DB 7 KHAVD 11
 RESULT 2
 ID O80978 PRELIMINARY; PRT; 117 AA.
 AC O80978;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE AT2G14730 protein (Hypothetical 13.6 kDa protein).
 GN AT2G14730.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.K., Bowman C.B., Barnstead M.E., Feldblum T.V.,
 RA Buel C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umam L.J., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*."
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX Town C.D., Kaul S.,
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC004705; AAC24184.1; --
 DR EMBL: AC005398; AAM15054.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 117 AA; 13650 MW; 1A3A5A9B684CC8BD CRC64;
 Query Match 100.0%; Score 27; DB 10; Length 117;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVD 5
 DB 17 KHAVD 21
 RESULT 3
 ID Q9SVF7 PRELIMINARY; PRT; 159 AA.
 AC Q9SVF7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Hypothetical 17.6 kDa protein.
 GN F22I13.50 OR AT4G38280.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Wedler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Schueller C.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Wedler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL035539; CAB37484.1; --
 DR EMBL: AL161593; CAB80493.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 159 AA; 17633 MW; 6A8683A7BB2E7C4B CRC64;
 Query Match 100.0%; Score 27; DB 10; Length 159;
 Best Local Similarity 100.0%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 DB 128 KHAVD 132

RESULT 4

Q99T24 PRELIMINARY; PRT; 179 AA.
 AC Q99T24; (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE Molybdenum cofactor biosynthesis protein.
 GN MOAB OR SC03179 OR SCE87.30C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Kienast H., Hopwood D.A.;
 RA "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]

Q99T24 PRELIMINARY; PRT; 199 AA.
 AC Q99T24; (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
 DE Hypothetical protein SAV1854.
 GN SAV1854 OR SAL1671.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=158878, 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE=213111952; PubMed=11418146;
 RA Kuroda M., Ochia T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanihisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003363; BAB58016.1; --
 DR EMBL; AP003363; BAB42940.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 199 AA; 22343 MW; 9FA58B2759B40F9D CRC64;

Query Match 100.0%; Score 27; DB 16; Length 199;
 Best Local Similarity 100.0%; Pred. No. 88; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0;

QY 1 KHAVD 5
 DB 55 KHAVD 59

RESULT 6

Q22147 PRELIMINARY; PRT; 211 AA.
 AC Q22147; (TREMELrel. 05, Created)
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE AT2G45250 protein.
 GN AT2G45250.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Rhee G., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umeyama L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrara A.J., Creasy T.H., Goodman H.M., Smerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;

Query Match 100.0%; Score 27; DB 16; Length 179;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 DB 158 KHAVD 162

RESULT 5

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC002387; AAB2645.1; --
 SQ SEQUENCE 211 AA; 23099 MW; EEB43B968BA3A513 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 211;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 Db 169 KHAVD 173

RESULT 7

Q9NEJ3 PRELIMINARY; PRT; 266 AA.
 ID Q9NEJ3
 AC Q9NEJ3
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Y116F11B.14 protein.
 GN Y116F11B.14.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sulston J.E., McLaughlin K.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AL132943; CAB81984.1; --
 SQ SEQUENCE 266 AA; 30175 MW; 63788ABDD0F41D43 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 266;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 Db 124 KHAVD 128

RESULT 8

Q9H315 PRELIMINARY; PRT; 274 AA.
 ID Q9H315
 AC Q9H315
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ARS protein.
 GN PNUTL2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20582451; PubMed=11146656;
 RA Larisch S., Yi Y., Lotan R., Kerner H., Eimerl S., Parks W.T.,
 RA Yossi G., Reffey S.B., de Caestecker M.P., Danielpour D.,
 RA Book-Melamed N., Timberg R., Duckett C., Lechneider R.J., Steller H.,
 RA Orly J., Kim S.-J., Roberts A.B.;
 RT "ARS, a novel mitochondrial septin-like protein mediates apoptosis
 RT dependent on its P-loop motif."
 RL Nat. Cell Biol. 2:915-921(2000).
 DR EMBL; AF176379; AAG45673.1; --

DR InterPro; IPR000038; GTP_Cell_Div.
 DR Pfam; PF00735; GTP_CDC; 1.
 DR ProDom; PD002565; GTP_Cell_Div; 1.
 SQ SEQUENCE 274 AA; 30809 MW; 7FA8D55492E9348F CRC64;

Query Match 100.0%; Score 27; DB 4; Length 274;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 Db 171 KHAVD 175

RESULT 9

O15988 PRELIMINARY; PRT; 293 AA.
 ID O15988
 AC O15988
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Troponin I.
 GN Troponin I.
 OS Patinopecten yessoensis (Ezo giant scallop) (Yesso scallop).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
 OC Pectinoidea; Pectinidae; Mizuhopecten.
 OC NCB1_TaxID=6573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hotta A.L.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB008006; BAA22853.1; --
 DR InterPro; IPR001978; Troponin.
 DR Pfam; PF00992; Troponin; 1.
 SQ SEQUENCE 293 AA; 34670 MW; 13EAAFC2DF651724 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 293;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 Db 280 KHAVD 284

RESULT 10

Q95XPI PRELIMINARY; PRT; 296 AA.
 ID Q95XPI
 AC Q95XPI
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 34.1 kDa protein.
 GN Y1G12B.22.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Du H., Lamar B., Wohldmann P.;
 RT "The sequence of C. elegans cosmid Y1G12B.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.


```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025726; AAK73923.1; -.
KW Hypothetical protein.
SQ SEQUENCE 296 AA; 34050 MW; 02C02888E06A3948 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 154 KHAVD 158

RESULT 11
O15987 PRELIMINARY; PRT; 314 AA.
AC O15987;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Troponin I.
OS Patinopecten yessoensis (Ezo giant scallop) (Yesso scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
OC Pectinoida; Pectinidae; Mizuhopecten.
OX NCBI_TaxID=6573;
RN [1]
RP SEQUENCE FROM N.A.
RA Hotta A.L.;
RT "Primary structures of the scallop, Patinopecten yessoensis, troponin
RT Is.";
DR EMBL; AB080805; BAA22852.1; -.
DR InterPro; IPR001978; Troponin.
DR Pfam; PF00392; Troponin; 1.
SQ SEQUENCE 314 AA; 37009 MW; 2FB61D644DBCAED8 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 314;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 297 KHAVD 301

RESULT 12
O95651 PRELIMINARY; PRT; 356 AA.
AC O95651;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Transcriptional activator GCN1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamagata K., Oda N., Furuta H., Vaxillaire M., Southam L., Boriraj V.,
RA Chen X., Oda Y., Takeda J., Yamada S., Nishigori H., Lebeau M.M.,
RA Lathrop M., Cox R.D., Bell G.I.;
RT "Transcription map of the 5' region surrounding the hepatocyte
RT nuclear factor-1a/MOBF3 gene on chromosome 12.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U98837; AAD00656.1; -.
FT NON_TER 1

```

```

FT NON_TER 356
SQ SEQUENCE 356 AA; 38828 MW; B1C80A13CA697006 CRC64;

Query Match 100.0%; Score 27; DB 4; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 156 KHAVD 160

RESULT 13
Q98BM3 PRELIMINARY; PRT; 415 AA.
AC Q98BM3;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Probable transglycolase.
GN MLR511.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MAPF303099; PubMed=11214969;
RX MEDLINE=21082930; PubMed=11214969;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawara K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51949.1; -.
KW Complete proteome.
SQ SEQUENCE 415 AA; 44644 MW; F4P2D66CDC56EA88 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 415;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 211 KHAVD 215

RESULT 14
Q8YEG7 PRELIMINARY; PRT; 427 AA.
AC Q8YEG7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Membrane-bound lytic murein transglycosylase B (EC 3.2.1.-).
GN BME11911.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mueh C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Renik G.,
RA Jablonowski L., Larsen N., D'Souza M., Berni A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen

```

```
RT Brucella melitensis";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009624; AAU53092.1; -.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 427 AA; 46638 MW; F427FCE95AAFCBE CRC64;

Query Match 100.0%; Score 27; DB 16; Length 427;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 223 KHAVD 227

RESULT 15
Q9K015 PRELIMINARY; PRT; 431 AA.
AC Q9K015;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Oxidoreductase, putative.
GN NMB0614.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
Hatt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002416; AAP1041.1; -.
DR TIGR; NMB0614; -.
DR InterPro; IPR000169; SHprot_acsite.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 431 AA; 48414 MW; 35D2FD1C22FAC6BD CRC64;

Query Match 100.0%; Score 27; DB 16; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 118 KHAVD 122

RESULT 16
Q9LOF7 PRELIMINARY; PRT; 473 AA.
AC Q9LOF7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative DEAD-box RNA helicase.
GN SCO4685 OR SCD31.10.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
Kinaishi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieser T., Larke L., Murphry L., Oliver K., O'Neill S.,
Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL; AL161803; CAB82053.1; -.
DR HSP; Q58083; IHV8.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
DR ATP-binding; Helicase.
KW ATP-binding; Helicase.
SQ SEQUENCE 473 AA; 50328 MW; 7A2CCBA86E1D175C CRC64;

Query Match 100.0%; Score 27; DB 16; Length 473;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 294 KHAVD 298

RESULT 17
Q930R6 PRELIMINARY; PRT; 481 AA.
AC Q930R6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RA0129.
GN RA0129 OR SMA0244.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
```

RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT Nucleotide sequence and predicted functions of the entire
RT Sinothrixobium mellotii pSymA megaplasmid.;
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AF007207; AAK64787.1; --
DR InterPro; IPR004113; FAD-oxidase_C;
DR InterPro; IPR001092; HLH_basic_
DR InterPro; IPR001575; Oxid_FAD_bind.
DR Pfam; PF02913; FAD-oxidase_C; 1.
DR Pfam; PF01565; FAD binding_4; 1.
DR PROSITE; PS00038; HELIX LOOP HELIX; UNKNOWN 1.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 481 AA; 51799 MW; A6C41069E70B91A CRC64;

Query Match 100.0%; Score 27; DB 16; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 464 KHAVD 468

RESULT 18
OS4116 PRELIMINARY; PRT; 498 AA.
ID 054116
AC 054116;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SC10A5.25c protein.
GN SC05920 OR SC10A5.25C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Kadenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdono-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).

DR EMBL; AL021529; CAA16457.1; --
DR HSSP; Q58083; 1HV8
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICc; 1.
DR PROSITE; PS00135; TRYPsin_SER; UNKNOWN 1.
KW ATP-binding; Helicase.
SQ SEQUENCE 498 AA; 52678 MW; 79F02BC9D5C96D10 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 498;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 318 KHAVD 322

RESULT 19
Q9JVD3 PRELIMINARY; PRT; 500 AA.
ID 09JVD3
AC 09JVD3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE N utilization substance protein A.
GN NMB1642.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Cirtone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Fizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AF002514; AAF41991.1; --
DR TIGR; NMB1642;
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR003029; S1_
DR Pfam; PF00013; KH-domain; 1.
DR Pfam; PF00575; S1; 1.
DR SMART; SM00322; KH; 1.
DR SMART; SM00316; S1; 1.
DR PROSITE; PS50084; KH_TYPE_1; 1.
KW Complete proteome.
SQ SEQUENCE 500 AA; 55751 MW; 753FAS0DD5F5B774 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 500;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
DB 310 KHAVD 314

RESULT 20
Q9JTB6

AD Q9JTB6 PRELIMINARY; PRT; 505 AA.
AC Q9JTB6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE N utilisation substance protein A.
GN NUSA OR NMA1896.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Davis K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
RL EMBL; AL162757; CAB85117.1; -.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR003029; SI.
DR Pfam; PF00013; KH-domain; 1.
DR Pfam; PF00575; SI; 1.
DR SMART; SM00322; KH; 2.
DR SMART; SM00316; SI; 1.
DR PROSITE; PS00084; KH_TYPE_1; 1.
KW Complete proteome.
SQ SEQUENCE 505 AA; 56418 MW; 5A0F080DCA99E5D7 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 505;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 315 KHAVD 319

RESULT 21
Q46243
ID Q46243 PRELIMINARY; PRT; 525 AA.
AC Q46243;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Nitrogenase.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93326641; PubMed=8334167;
RX Zinoni F., Robson R.M., Robson R.L.;
RT "Organization of potential alternative nitrogenase genes from
RT Clostridium pasteurianum."
RL Biochim. Biophys. Acta 1174:83-86(1993).
DR EMBL; L09762; AAC36970.1; -.
DR HSP; P00467; IMIO.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR000318; Nitrogenase compl.
DR InterPro; IPR000510; Oxred_nitrogenase.
DR Pfam; PF00148; oxidored_nitro; 1.
DR TIGRfam; TIGR01284; alt_nitrog_alph; 1.
DR PROSITE; PS00225; CRYSTALLIN BETAGAMMA; UNKNOWN_1.
DR PROSITE; PS00699; NITROGENASE_1; 1.
DR PROSITE; PS00090; NITROGENASE_1_2; 1.

SQ SEQUENCE 525 AA; 59283 MW; A70F5906FEDB0D73 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 17 KHAVD 21

RESULT 22
Q9XUR9
ID Q9XUR9 PRELIMINARY; PRT; 534 AA.
AC Q9XUR9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE R10E8.4 protein.
GN R10E8.4.
OS Caenorhabditis elegans.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Ainscough R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z81576; CAB04641.1; -.
SQ SEQUENCE 534 AA; 60700 MW; A33FA1066CF75071 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 534;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVD 5
Db 419 KHAVD 423

RESULT 23
Q9SOV3
ID Q9SOV3 PRELIMINARY; PRT; 535 AA.
AC Q9SOV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Dinitrogenase alpha subunit.
GN ANFD.
OS Nitrogen fixing bacterium ANFK33.
OC Bacteria.
OX NCBI_TaxID=96027;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20011232; PubMed=10543806;
RX Noda S., Ohkuma M., Usami R., Horikoshi K., Kudo T.;
RT "Culture-independent characterization of gene responsible for nitrogen
RT fixation in the symbiotic microbial community in the gut of the
RT termite Neotermes koshunensis."
RL Appl. Environ. Microbiol. 65:4935-4942(1999).
DR EMBL; AB027751; BAA86281.1; -.
DR HSP; P00467; IMIO.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR000318; Nitrogenase compl.
DR InterPro; IPR000510; Oxred_nitrogenase.
DR Pfam; PF00148; oxidored_nitro; 1.

DR TIGRPFAMs; TIGR01284; alt_nitrog_alph; 1.
 DR PROSITE; PS00225; CRYSTALLIN BETAGAMMA; UNKNOWN_1.
 DR PROSITE; PS00699; NITROGENASE_1_1; 1.
 DR PROSITE; PS00090; NITROGENASE_1_2; 1.
 SQ SEQUENCE 535 AA; 60478 MW; EB8A29339AEDCFPE CRC64;

Query Match 100.0%; Score 27; DB 2; Length 535;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 |||||
 DB 17 KHAVD 21

RESULT 24

ID Q9ZBJ9 PRELIMINARY; PRT; 601 AA.
 AC Q9ZBJ9;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Putative oxidoreductase.
 GN SC06475 OR SC9C7.11C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrrell B.G., Rajandream M.A.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8943436;
 RA Redenbach M., Kieser H.M., Denaparte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitch B., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)";
 RL Nature 417:141-147(2002).
 DR EMBL; AL035163; CAA22723.1; --
 DR HSSP; P00348; 3HDH.
 DR InterPro; IPR002135; 3HCDH.
 DR InterPro; IPR002025; NAD_binding.
 DR Pfam; PF00725; 3HCDH; 2.
 DR Pfam; PF02737; 3HCDH; N: 2.
 SQ SEQUENCE 601 AA; 62855 MW; 61BEE1F421103B54 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 601;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 |||||
 DB 492 KHAVD 496

RESULT 25

ID Q54907 PRELIMINARY; PRT; 656 AA.
 AC Q54907;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Oligopeptidasepermease.
 GN OPPA.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SEROTYPE M49;
 RX MEDLINE=97039725; PubMed=885277;
 RA Podbielski A., Pohl B., Woischnik M., Koerner C., Schmidt K.H.,
 RA Rodzinski E., Leonard B.A.B.;
 RT "Molecular characterization of a group A streptococcal (GAS)
 oligopeptidase permease (Opp) and its effect on cysteine protease
 production";
 RL Mol. Microbiol. 21:1087-1099(1996).
 DR EMBL; X89237; CAA61522.1; --
 DR InterPro; IPR000914; SBP_bac_5.
 DR Pfam; PF00496; SBP_bac_5; 1.
 DR PROSITE; PS01040; SBP_BACTERIAL_5; UNKNOWN_1.
 SQ SEQUENCE 656 AA; 72026 MW; 60BB5EE0859439A2 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 656;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVD 5
 |||||
 DB 127 KHAVD 131

Search completed: June 20, 2003, 20:56:56
 Job time : 29.3636 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 31.9318 Seconds
(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-14

Perfect score: 27

Sequence: 1 DRAVK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

A_Geneseq_101002.*

1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27	100.0	5	20	AAV17155
2	27	100.0	5	21	AAV73804
3	27	100.0	5	22	AA665374
4	27	100.0	5	23	AA47481
5	27	100.0	103	23	ABP35003
6	27	100.0	140	21	AA41191
7	27	100.0	489	22	AAU56005
8	27	100.0	508	18	AAW89793
9	27	100.0	510	22	AA691315
10	27	100.0	510	22	AA676687

RESULT 1

AAV17155

ID AAV17155 standard; peptide; 5 AA.

ALIGNMENTS

11	27	100.0	885	22	ABB61554
12	27	100.0	1018	21	AA07773
13	27	100.0	1018	22	AA05043
14	27	100.0	1066	22	AA07918
15	27	100.0	1066	22	AA080134
16	27	100.0	1113	22	AA091525
17	27	100.0	1113	22	AA07917
18	27	100.0	1113	22	AA080133
19	27	100.0	1113	23	AA079297
20	27	100.0	1245	22	ABG03370
21	27	100.0	1277	22	ABG14366
22	27	100.0	2248	22	ABG14366
23	26	96.3	167	20	AAW94614
24	26	96.3	167	20	AAW89864
25	26	96.3	167	21	AA080189
26	26	96.3	238	22	AA077339
27	26	96.3	279	23	AB084768
28	26	96.3	292	22	AA060860
29	26	96.3	297	23	ABP28408
30	26	96.3	465	22	AA096333
31	26	96.3	509	22	AA093914
32	26	96.3	509	22	AA093694
33	26	96.3	511	22	AA041100
34	26	96.3	515	22	AAU34315
35	26	96.3	525	22	AAU37333
36	26	96.3	750	22	ABB71433
37	26	96.3	915	22	ABB70955
38	26	96.3	1394	22	ABB61170
39	26	96.3	2492	23	AAE18302
40	24	88.9	43	22	AA064956
41	24	88.9	43	22	AA037892
42	24	88.9	68	21	AAI55658
43	24	88.9	83	22	AA000063
44	24	88.9	85	21	AA026793
45	24	88.9	88	19	AAW98426
46	24	88.9	92	21	AA053653
47	24	88.9	107	23	ABP32045
48	24	88.9	113	23	ABP01911
49	24	88.9	116	22	AAU53148
50	24	88.9	122	21	AA02650
51	24	88.9	143	20	AA038386
52	24	88.9	148	21	AA027136
53	24	88.9	149	20	AA02624
54	24	88.9	153	21	AA028871
55	24	88.9	155	23	ABG60170
56	24	88.9	161	21	AA012580
57	24	88.9	162	21	AA044619
58	24	88.9	175	22	AA059821
59	24	88.9	186	20	AA02622
60	24	88.9	186	21	AAI57948
61	24	88.9	191	22	AAU45172
62	24	88.9	193	15	AA057059
63	24	88.9	193	15	AA057060
64	24	88.9	193	21	AA097059
65	24	88.9	195	22	AA05832
66	24	88.9	217	22	ABG01051
67	24	88.9	217	22	ABG18817
68	24	88.9	217	22	ABG23035
69	24	88.9	218	23	AAU97298
70	24	88.9	236	23	AB054768
71	24	88.9	280	21	AA033848
72	24	88.9	281	23	ABP30321
73	24	88.9	285	21	AA033847
74	24	88.9	294	23	ABP28331
75	24	88.9	299	16	AA075416

Drosophila melanog
Large subunit of c
B. lactofermentum
Corynebacterium gl
Corynebacterium gl
C glutamicum prote
Corynebacterium gl
Corynebacterium gl
Polyptide with c
Novel human diagno
Novel human diagno
Drosophila melanog
Helicobacter pylor
Antigen 1 from clu
Amino acid sequenc
Novel bone marrow
DNA polymerase III
Mevalonate pathway
Streptococcus poly
Pucative P. abyssi
Human polyptide
Human protein sequ
Human polyptide
Staphylococcus aur
Staphylococcus aur
Drosophila melanog
Drosophila melanog
Drosophila melanog
Venezuelan equine
Human brain expres
Peptide #11929 enc
H. pylori ycfB pep
Cytosolic pyruvate
Zea mays protein f
H. pylori GHP0 384
Arabidopsis thalia
Human RNA polymera
Human ORF protein
Propionibacterium
Human secreted pro
Human secreted pro
Zea mays protein f
A differentially s
Zea mays protein f
Human DITHP polype
Zea mays protein f
Zea mays protein f
TutE protein #5
A human TRAM-1 pro
Human transmembran
Propionibacterium
P. aeruginosa aligu
P. aeruginosa aligu
P. aeruginosa aligu
Human protein sequ
Novel human diagno
Novel human diagno
Novel human diagno
Agrobacterium tume
Lactococcus lactis
Arabidopsis thalia
Streptococcus poly
Arabidopsis thalia
Streptococcus poly
Rat regucalcin, a

XX AAY17155;
 XX 20-SEP-1999 (first entry)
 XX
 XX Cadherin-mediated adhesion modulating cyclic peptide.
 DE
 DE Cell adhesion modulation; CAM; synaptic stability; cadherin;
 KW cadherin-mediated adhesion; drug delivery; cell adhesion; tumour;
 KW wound healing; neurite outgrowth; cyclic; circular.
 XX
 XX Synthetic.
 XX WO9933875-A1.
 PN
 XX 08-JUL-1999.
 PD
 XX
 XX 23-DEC-1998; 98WO-CA01207.
 PF
 XX 23-DEC-1997; 97US-0996679.
 PR
 XX (UYMC-) UNIV MCGILL.
 PA
 XX Blaschuk OW, Gour BJ;
 PI
 XX WPI; 1999-430231/36.
 DR
 XX
 XX Cyclic peptide cell adhesion modulating agents, useful for
 PT modulating synaptic stability
 PT
 XX
 PS Claim 17; Page 65; 144pp; English.
 XX
 XX The invention provides cyclic peptide cell adhesion modulating (CAM)
 CC agents that comprises a His-Ala-Val recognition sequence. Also provided
 CC is a method for inhibiting synaptic stability in a mammal that comprises
 CC administering to a mammal a therapeutically effective amount of a CAM
 CC agent that inhibits cadherin-mediated adhesion, where the agent comprises
 CC a cyclic peptide having a peptide ring, and where the sequence His-Ala-
 CC Val is present within the peptide ring. The cyclic peptides are cell
 CC adhesion modulating agents that inhibit cadherin-mediated adhesion. They
 CC can be used in a method for inhibiting synaptic stability in mammals. The
 CC agents can be used to treat diseases or other conditions characterized by
 CC undesirable cell adhesion or to facilitate drug delivery to a specific
 CC tissue or tumour. Alternatively the agents may be used to enhance cell
 CC adhesion (e.g. to supplement or replace stitches or to facilitate wound
 CC healing) or to enhance or direct neurite outgrowth.
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 27; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 Db |||||
 1 DHAVK 5
 RESULT 2
 AAY73804
 ID AAY73804 standard; Peptide; 5 AA.
 XX
 AC AAY73804;
 XX
 XX 07-MAR-2000 (first entry)
 DT
 XX N-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:310.
 DE
 XX Diagnosis; cancer; OB-cadherin; N-cadherin; E-cadherin; CAR; cyclic;
 KW cell adhesion recognition sequence; leukaemia; metastasis;
 KW prostate cancer; breast cancer; ovarian cancer; carcinoma.
 XX
 OS Synthetic.

OS Homo sapiens.
 XX WO9957565-A2.
 PN
 XX 11-NOV-1999.
 PD
 XX
 XX 05-MAY-1999; 99WO-CA00362.
 PF
 XX 05-MAY-1998; 98US-0073040.
 PR
 XX 06-NOV-1998; 98US-0187859.
 PR
 XX 20-JAN-1999; 99US-0234395.
 PA
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX Blaschuk OW, Gour BJ, Byers S;
 PI
 XX WPI; 2000-062165/05.
 DR
 XX
 XX Detecting expression of OB-cadherin or N-cadherin used for diagnosing
 PT and evaluating cancer
 PT
 XX
 PS Disclosure; Page 22; 56pp; English.
 XX
 XX The present invention describes methods which have been developed for
 CC detecting and evaluating cancer that are based on the finding that
 CC OB-cadherin and N-cadherin are expressed by metastatic carcinoma cells,
 CC but not by highly differentiated, poorly invasive carcinomas. A method
 CC for determining the presence or absence of a cancer in a patient
 CC comprises: (1) contacting a biological sample from the patient with a
 CC binding agent that specifically binds to OB- or N-cadherin OR with an
 CC oligonucleotide that hybridises to a polynucleotide that encodes OB- or
 CC N-cadherin; and (2) detecting in the sample an amount of polypeptide
 CC that binds to the binding agent OR the amount of polynucleotide that
 CC hybridises to the oligonucleotide, relative to a predetermined cut-off
 CC value, and determining the presence or absence of cancer in the patient
 CC from this. The methods from the present invention can be used to
 CC determine the metastatic potential of a cancer. The methods may be used
 CC to detect a metastatic cancer in a patient, to monitor progression of a
 CC cancer, or to evaluate the metastatic potential of a cancer. Cancers
 CC which may be evaluated using the methods include leukemia, prostate
 CC cancer, breast cancer and ovarian cancer. AAY73501 to AAY73813 represent
 CC peptide sequences used in the exemplification of the present invention.
 CC Specifically, AAY73503 to AAY73789 represent OB-cadherin cell adhesion
 CC recognition (CAR) peptides and AAY73790 to AAY73808 represent N-cadherin
 CC CAR peptides. AAY73906 to AAY73915 represent PCR primers used in examples
 CC from the present invention.
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 27; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 Db |||||
 1 DHAVK 5
 RESULT 3
 AAG65374
 ID AAG65374 standard; peptide; 5 AA.
 XX
 AC AAG65374;
 XX
 XX 30-NOV-2001 (first entry)
 DT
 XX
 XX Cyclic peptide with cadherin CAR sequence.
 DE
 XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 KW cytostatic; vulnery; immunomodulator; vasotrophic; neuroprotective;
 KW cerebroprotective; muscular; cyclic.
 XX
 OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal acetylation"
 FT Modified-site 5
 FT /note= "C-terminal amide"
 XX XX
 PN WO200153331-A2.
 XX
 PD 26-JUL-2001.
 XX
 XX 24-JAN-2001; 2001WO-US02508.
 PF
 XX 24-JAN-2000; 2000US-0491078.
 PR
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 PA
 XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S,
 PI Hu Z;
 XX WPI; 2001-549899/61.
 DR
 XX Cell adhesion modulating agent used for enhancing delivery of drug to
 PT tumor comprises imidazole compounds -
 FT
 XX Examples; Fig 3A-I; 436pp; English.
 PS
 XX The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cyclic peptide with
 CC cadherin cell adhesion recognition (CAR) sequence.
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 27; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 DB 1 DHAVK 5
 RESULT 4
 AAM47481
 ID AAM47481 standard; peptide; 5 AA.
 XX
 AC AAM47481;
 XX
 DT 12-FEB-2002 (first entry)
 DE
 DE Cyclic peptide endothelial cell adhesion modulator #3.
 XX
 XX Cadherin; cytostatic; gynecological; endometriosis;
 KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 KW cyclic.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Linked to residue 5 to form a cyclic peptide,
 FT N-terminal acetyl or N-terminal CH3-SO2 group"
 FT

FT Modified-site 5
 FT /note= "Linked to residue 1 to form a cyclic peptide,
 FT C-terminal amide"
 XX
 PN WO200177146-A2.
 XX
 PD 18-OCT-2001.
 XX
 XX 09-APR-2001; 2001WO-US11669.
 PF
 XX 07-APR-2000; 2000US-0544782.
 PR
 XX (UYMC-) UNIV MCGILL.
 PA
 XX Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 PI WPI; 2002-049129/06.
 DR
 XX Modulating endothelial cell adhesion for inhibiting development of
 PT endometriosis, increasing blood flow to tumor in a mammal, by
 PT contacting cell with a cyclic peptide having cadherin cell adhesion
 PT recognition sequence -
 XX
 PS Claim 6; Page 64; 139pp; English.
 XX
 CC The present invention relates to a method for modulating endothelial cell
 CC adhesion. The method comprises contacting an endothelial cell with a
 CC modulating peptide comprising a cadherin cell adhesion recognition
 CC sequence (this Ala Val) within a cyclic peptide ring. The method is useful
 CC for inhibiting angiogenesis, increasing vasopermeability, increasing the
 CC blood flow to a tumour, disrupting neovasculation and inhibiting the
 CC development of endometriosis in a mammal. The modulating peptide reduces
 CC unwanted endothelial adhesion occurring between tumour cells, tumour
 CC cells and normal cells, normal cells as a result of surgery, injury,
 CC chemotherapy, disease and inflammation. The present sequence is one such
 CC modulating peptide.
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 27; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 DB 1 DHAVK 5
 RESULT 5
 ABP35003
 ID ABP35003 standard; Protein; 103 AA.
 XX
 AC ABP35003;
 XX
 DT 09-JUL-2002 (first entry)
 DE
 DE Human kinase-like ORP3976 protein, SEQ ID NO:7952.
 XX
 XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulvare;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.

XX PN W0200190366-A2.
XX PD 29-NOV-2001.
XX XX
XX PF 24-MAY-2001; 2001WO-US17076.
XX PR 24-MAY-2000; 2000US-206690P.
XX PA (CURA-) CURAGEN CORP.
XX PI Leach MD, Shinkets RA;
XX DR WFI; 2002-106200/14.
XX DR N-PSDB; ABN79029.
XX XX
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
XX PT preventing and treating cardiovascular disease, neurodegenerative,
XX PT hyperproliferative disorders and disorders related to organ
XX PT transplantation
XX PS Claim 10; Page 2234; 2508pp; English.
XX CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
XX CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
XX CC ABN79587 represent cDNAs encoding them. The invention also encompasses
XX CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
XX CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
XX CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
XX CC polynucleotides, the recombinant production of ORFX proteins, antibodies
XX CC specific for ORFX proteins, methods for detecting ORFX polynucleotides and
XX CC polypeptides, methods of screening for modulators of ORFX expression or
XX CC activity, and methods of screening individuals for a predisposition to an
XX CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
XX CC range of biological activities, such as cytokine, cell proliferation,
XX CC cell differentiation, immune modulation, haematopoiesis regulation,
XX CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
XX CC chemokinetic activity, haemostatic activity, thrombolytic activity,
XX CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
XX CC and antifertility activity, and may also be involved in the determination
XX CC of bodily characteristics, fertility and behaviour. ORFX proteins,
XX CC nucleic acids and antibodies may be used in the treatment of cancers,
XX CC other proliferative disorders such as psoriasis and benign tumours,
XX CC neurological disorders such as epilepsy and Alzheimer's disease,
XX CC cardiovascular diseases, immune system disorders, disorders related to
XX CC organ transplantation, disorders of tissue growth and regeneration,
XX CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
XX CC storage disease, and infectious diseases caused by viral, bacterial,
XX CC fungal and other pathogens. ORFX nucleic acids may also be used as a
XX CC source of primers and probes, in the detection of ORFX genomic sequences
XX CC or transcripts, in the identification and cloning of homologous
XX CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
XX CC nucleic acids may additionally be used to produce transgenic animals
XX CC which may be useful for studying the function and/or activity of ORFX
XX CC protein, and in drug screening. The ORFX proteins may also be used as
XX CC immunogens to generate specific antibodies, which are useful in the
XX CC diagnosis, treatment and monitoring of ORFX-associated diseases.
XX SQ Sequence 103 AA;

Query Match 100.0%; Score 27; DB 23; Length 103;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
Db 83 DHAVK 87

RESULT 6
AAG41191
ID AAG41191 standard; Protein; 140 AA.
XX

AC AAG41191;
XX 18-OCT-2000 (first entry)
XX XX
XX DE Zea mays protein fragment SEQ ID NO: 51218.
XX XX
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence; corn.
XX OS Zea mays subsp. mays.
XX XX
XX PN EP1033405-A2.
XX XX
XX PD 06-SEP-2000.
XX XX
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 29-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 05-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132563.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 16-JUN-1999; 99US-0139453.
XX PR 17-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 24-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149369.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150894.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160880.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 27; DB 21; Length 140;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DHAVK 5
 Db 20 DHAVK 24

RESULT 7

AAU56005
 ID AAU56005 standard; Protein; 489 AA
 XX
 AC AAU56005;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #16901.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

XX 01-NOV-2001.

PD 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59573.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX Example 1; SEQ ID No 17200; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 489 AA;

Query Match 100.0%; Score 27; DB 22; Length 489;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

DB 381 DHAVK 385

RESULT 8

AAW89793

ID AAW89793 standard; Protein; 508 AA.

XX AAW89793;

XX 16-MAR-1999 (first entry)

XX Staphylococcus aureus protein SEQ ID #5241.

XX Computer readable medium; vaccine; S. aureus infection; immunodetection;

PN EP1108790-A2.

KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome.

OS Staphylococcus aureus.

PN EP786519-A2.

XX 30-JUL-1997.

PD 07-JAN-1997; 97EP-0100117.

XX 05-JAN-1996; 96US-0009861.

PR (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, Choi GH, Dillon FJ, Fannon MR, Kunesch CA;

PI Rosen CA;

XX WPI; 1997-374922/35.

DR Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S. aureus vaccines

XX Claim 23; Page 3242-3244; 3271pp; English.

XX This sequence represents a Staphylococcus aureus protein sequence of the
 CC invention. The DNA sequences encoding the S. aureus proteins are recorded
 CC on a computer readable medium, preferably selected from a floppy or hard
 CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
 CC Homology searches using the S. aureus DNA sequences allows putative
 CC functions to be assigned so that protein-encoding or regulatory regions
 CC of commercial, therapeutic or industrial importance can be obtained.
 CC Specifically, sequences which are likely to encode antigens have been
 CC identified and these polypeptides can be used in a vaccine composition
 CC against S. aureus infection. The polypeptides can also be used in a kit
 CC for the immunodetection of S. aureus in a sample. S. aureus is implicated
 CC in numerous human diseases, including cellulitis, eyelid infections, food
 CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
 CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
 CC DNA sequences can be used for recombinant production of the polypeptides.
 CC The new DNA sequences (and their fragments) are useful as primers or
 CC probes for isolating homologues of any of the 5191 S. aureus DNA sequences
 CC contained on the computer readable medium.

XX SQ Sequence 508 AA;

Query Match 100.0%; Score 27; DB 18; Length 508;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

DB 66 DHAVK 70

RESULT 9

AAAG91315

ID AAG91315 standard; Protein; 510 AA.

XX AAG91315;

XX 26-SEP-2001 (first entry)

DT C glutamicum protein fragment SEQ ID NO: 5069.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

XX Corynebacterium glutamicum.

XX EP1108790-A2.

XX PD 20-JUN-2001.
 XX PF 18-DEC-2000; 2000EP-0127588.
 XX PR 16-DEC-1999; 99JP-0377484.
 XX PR 07-APR-2000; 2000JP-0159162.
 XX PR 03-AUG-2000; 2000JP-0280988.
 XX PA (KYOWA) KYOWA HAKKO KOGYO KK.
 XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX DR WPI; 2001-376931/40.
 XX DR N-PSDB; AAH66534.
 XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 XX PT mutation point of a gene, measuring expression of a gene, analysing
 XX PT expression profile or pattern of a gene and identifying homologous gene
 XX PS Claim 17; SEQ ID NO: 5069; 246pp + Sequence Listing; English.
 XX CC The present invention provides a number of nucleotide and protein
 XX CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 XX CC are useful for identifying the mutation point of a gene derived from a
 XX CC mutant of coryneform bacterium, measuring expression amount and
 XX CC analysing the expression profile or expression pattern of a gene derived
 XX CC from Coryneform bacterium, and identifying a homologue of a gene derived
 XX CC from Coryneform bacterium. Coryneform bacteria are useful for producing
 XX CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 XX CC particularly L-lysine. The present sequence is a protein described
 XX CC in the exemplification of the invention.
 XX CC Note: The sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic format directly from the
 XX CC European Patent Office.
 XX SQ Sequence 510 AA;
 Query Match 100.0%; Score 27; DB 22; Length 510;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 DB 418 DHAVK 422
 RESULT 10
 AAB76687
 ID AAB76687 standard; Protein; 510 AA.
 XX AC AAB76687;
 XX DT 11-APR-2001 (first entry)
 XX DE Corynebacterium glutamicum MCT protein SEQ ID NO:356.
 XX KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
 XX KW membrane construction and membrane transport protein; petroleum spill;
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
 KW identification; microorganism; fine chemical production; transformation;
 KW genome mapping; genetic engineering.
 XX OS Corynebacterium glutamicum.
 XX PN W0200100805-A2.
 XX PD 04-JAN-2001.
 XX PF 23-JUN-2000; 2000WO-IB00926.
 XX

25-JUN-1999; 99US-0141031.
 08-JUL-1999; 99DE-1031454.
 08-JUL-1999; 99DE-1031478.
 08-JUL-1999; 99DE-1031563.
 09-JUL-1999; 99DE-1032122.
 09-JUL-1999; 99DE-1032124.
 09-JUL-1999; 99DE-1032125.
 09-JUL-1999; 99DE-1032128.
 09-JUL-1999; 99DE-1032180.
 09-JUL-1999; 99DE-1032182.
 09-JUL-1999; 99DE-1032190.
 09-JUL-1999; 99DE-1032191.
 09-JUL-1999; 99DE-1032209.
 09-JUL-1999; 99DE-1032212.
 09-JUL-1999; 99DE-1032227.
 09-JUL-1999; 99DE-1032228.
 09-JUL-1999; 99DE-1032229.
 09-JUL-1999; 99DE-1032230.
 14-JUL-1999; 99DE-1032927.
 14-JUL-1999; 99DE-1033005.
 14-JUL-1999; 99DE-1033006.
 27-AUG-1999; 99DE-1040764.
 27-AUG-1999; 99DE-1040765.
 27-AUG-1999; 99DE-1040766.
 27-AUG-1999; 99DE-1040830.
 27-AUG-1999; 99DE-1040831.
 27-AUG-1999; 99DE-1040832.
 27-AUG-1999; 99DE-1040833.
 31-AUG-1999; 99DE-1041378.
 31-AUG-1999; 99DE-1041379.
 31-AUG-1999; 99DE-1041395.
 03-SEP-1999; 99DE-1042077.
 03-SEP-1999; 99DE-1042078.
 03-SEP-1999; 99DE-1042079.
 03-SEP-1999; 99DE-1042088.
 (BADI) BASF AG.
 Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 WPI; 2001-071486/08.
 N-PSDB; AAF67920.
 Corynebacterium glutamicum nucleic acids encoding membrane construction
 and membrane transport proteins or their portions, useful for typing or
 identifying C. glutamicum or related bacteria, and as markers for
 transformation.
 Claim 20; Page 682-683; 1119pp; English.
 AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
 construction and membrane transport (MCT) proteins given in AAB76510 to
 AAB76847. The MCT nucleic acids and proteins are useful in the
 identification of microorganisms which can be used to produce fine
 chemicals, for modulating fine chemical production in C. glutamicum or
 related bacteria (e.g. Brevibacterium lactofermentum), the typing or
 identification of C. glutamicum or related bacteria, as reference points
 for mapping C. glutamicum genome, and as markers for transformation.
 AAF68082 and AAF68082 represent sequencing primers which are used in an
 example from the present invention.
 Sequence 510 AA;
 Query Match 100.0%; Score 27; DB 22; Length 510;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 DB 418 DHAVK 422
 RESULT 11

ABB61554
ID ABB61554 standard; Protein; 885 AA.
XX AC ABB61554;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 11454.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX WO20017042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL05657.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX PS Disclosure; SEQ ID NO 11454; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABBS7737-ABBS72072).
XX CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 885 AA;
XX
Query Match 100.0%; Score 27; DB 22; Length 885;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
DB 206 DHAVK 210
RESULT 12
AAB07773
ID AAB07773 standard; Protein; 1018 AA.
XX AC AAB07773;
XX DT 07-NOV-2000 (first entry)
XX DE Large subunit of the carbamoyl-phosphate synthetase polypeptide.
XX KW Carbamoyl-phosphate synthetase; carA gene; carB gene; carbonyl phosphate;
XX carbonic acid; citrulline; ornithine; L-arginine; coryneform bacterium.
XX OS Brevibacterium lactofermentum.
XX
Query Match 100.0%; Score 27; DB 21; Length 1018;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
DB 1013 DHAVK 1017
RESULT 13
AAB50543
ID AAB50543 standard; Protein; 1018 AA.
XX AC AAB50543;
XX DT 20-MAR-2001 (first entry)
XX DE B. lactofermentum pyrimidine biosynthetic system enzyme protein #4.
XX KW Brevibacterium lactofermentum; coryneform; microbe; dihydroorotase;
XX pyrimidine biosynthetic system; aspartate carbamoyltransferase;
XX orotidylate decarboxylase.
XX OS Brevibacterium lactofermentum.
XX PN JP2000287688-A.
XX PD 17-OCT-2000.
XX PF 08-APR-1999; 99JP-0100989.
XX PR 08-APR-1999; 99JP-0100989.
XX PA (AJIN) AJINOMOTO KK.
XX WPI; 2001-065574/08.
XX DR N-PSDB; AAC90812.
XX PT A pyrimidine biosynthetic system enzyme gene of a coryneform microbe
XX

XX EP1026247-A1.
XX PD 09-AUG-2000.
XX PF 01-FEB-2000; 2000EP-0101648.
XX PR 01-FEB-1999; 99JP-0024149.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Kuwabara Y, Hashiguchi K, Nakamatsu T, Kurahashi O, Mori Y, Ito H;
XX WPI; 2000-500253/45.
XX DR N-PSDB; AAS9353.
XX PT Novel carbamoyl-phosphate synthetase protein and gene obtained from
XX coryneform bacterium useful for transforming microorganisms for
XX producing L-arginine -
XX PS Claim 8; Page 21-24; 31pp; English.
XX CC AAB07772-73 represent small and large subunits of the small and large
XX subunits of a carbamoyl-phosphate synthetase (CPS) polypeptide. They are
XX encoded by the carA and carB genes. CPS is an enzyme that catalyses the
XX reactions producing carbonyl phosphate from carbonic acid. The carbonyl
XX phosphate thus produces serves as a source of carbonyl group required
XX for the reaction producing citrulline from ornithine in the L-arginine
XX biosynthetic pathway. A coryneform bacterium transformed with the CPS
XX polynucleotide is useful for producing L-arginine.
XX SQ Sequence 1018 AA;
XX
Query Match 100.0%; Score 27; DB 21; Length 1018;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
DB 1013 DHAVK 1017
RESULT 13
AAB50543
ID AAB50543 standard; Protein; 1018 AA.
XX AC AAB50543;
XX DT 20-MAR-2001 (first entry)
XX DE B. lactofermentum pyrimidine biosynthetic system enzyme protein #4.
XX KW Brevibacterium lactofermentum; coryneform; microbe; dihydroorotase;
XX pyrimidine biosynthetic system; aspartate carbamoyltransferase;
XX orotidylate decarboxylase.
XX OS Brevibacterium lactofermentum.
XX PN JP2000287688-A.
XX PD 17-OCT-2000.
XX PF 08-APR-1999; 99JP-0100989.
XX PR 08-APR-1999; 99JP-0100989.
XX PA (AJIN) AJINOMOTO KK.
XX WPI; 2001-065574/08.
XX DR N-PSDB; AAC90812.
XX PT A pyrimidine biosynthetic system enzyme gene of a coryneform microbe
XX

PS Claim 8; Page 21-23; 25pp; Japanese.
 XX The present invention describes a pyrimidine biosynthetic system enzyme
 CC gene of a corynebacterium microbe e.g. Brevibacterium lactofermentum. Also
 CC described are: (1) a DNA sequence encoding a protein comprising a
 CC defined 298 amino acid sequence (P1) (AAB50540), or a sequence in which
 CC at least one amino acid residue is replaced, deleted, inserted, added or
 CC reversed in P1, where the protein has aspartate carbamoyltransferase
 CC activity; (2) a DNA encoding a protein comprising a defined 397 amino
 CC acid sequence (P2) (AAB50541) or a sequence in which at least one amino
 CC acid residue is replaced, deleted, inserted, added or reversed in P2,
 CC where the protein has dihydroorotase activity; (3) a DNA encoding a
 CC protein comprising a defined 278 amino acid sequence (P3) (AAB50548) or a
 CC sequence in which at least one amino acid residue is replaced, deleted,
 CC inserted, added or reversed in P3, where the protein has orotidylate
 CC decarboxylase activity; (4) a corynebacterium microbe transformed by the DNA;
 CC and (5) the proteins P1-P3 and their variants encoding by the DNA. The
 CC gene can be used for breeding nucleic acid producing microbes. The
 CC present sequence represents a protein encoded by a chromosome DNA
 CC fragment from Brevibacterium lactofermentum (AAC90812), which is a
 CC pyrimidine biosynthetic system enzyme.
 XX SQ Sequence 1018 AA;
 Query Match 100.0%; Score 27; DB 22; Length 1018;
 Best Local Similarity 100.0%; Pred No. 5, 7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 Db 1013 DHAVK 1017
 RESULT 14
 AAB79818
 ID AAB79818 standard; Protein; 1066 AA.
 AC AAB79818;
 XX AAB79818;
 DT 30-APR-2001 (first entry)
 DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:370.
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
 KW fine chemical production; microorganism; organic acid; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
 XX OS Corynebacterium glutamicum.
 XX WO200100843-A2.
 PN 04-JAN-2001.
 PD 23-JUN-2000; 2000WO-IB00923.
 PF 25-JUN-1999; 99US-0141031.
 XX 01-JUL-1999; 99DE-1030476.
 PR 02-JUL-1999; 99US-0142101.
 PR 08-JUL-1999; 99DE-1031415.
 PR 08-JUL-1999; 99DE-1031418.
 PR 08-JUL-1999; 99DE-1031419.
 PR 08-JUL-1999; 99DE-1031420.
 PR 08-JUL-1999; 99DE-1031424.
 PR 08-JUL-1999; 99DE-1031428.
 PR 08-JUL-1999; 99DE-1031434.
 PR 08-JUL-1999; 99DE-1031435.
 PR 08-JUL-1999; 99DE-1031443.
 PR 08-JUL-1999; 99DE-1031453.
 PR 08-JUL-1999; 99DE-1031457.
 PR 08-JUL-1999; 99DE-1031465.
 PR 08-JUL-1999; 99DE-1031478.

PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031541.
 PR 08-JUL-1999; 99DE-1031573.
 PR 08-JUL-1999; 99DE-1031592.
 PR 08-JUL-1999; 99DE-1031632.
 PR 08-JUL-1999; 99DE-1031634.
 PR 08-JUL-1999; 99DE-1031636.
 PR 09-JUL-1999; 99DE-1032125.
 PR 09-JUL-1999; 99DE-1032126.
 PR 09-JUL-1999; 99DE-1032130.
 PR 09-JUL-1999; 99DE-1032186.
 PR 09-JUL-1999; 99DE-1032206.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032228.
 PR 09-JUL-1999; 99DE-1032229.
 PR 09-JUL-1999; 99DE-1032230.
 PR 14-JUL-1999; 99DE-1032922.
 PR 14-JUL-1999; 99DE-1032926.
 PR 14-JUL-1999; 99DE-1032928.
 PR 14-JUL-1999; 99DE-1033004.
 PR 14-JUL-1999; 99DE-1033005.
 PR 14-JUL-1999; 99DE-1033006.
 PR 12-AUG-1999; 99US-0148613.
 PR 27-AUG-1999; 99DE-1040764.
 PR 27-AUG-1999; 99DE-1040765.
 PR 27-AUG-1999; 99DE-1040766.
 PR 27-AUG-1999; 99DE-1040832.
 PR 31-AUG-1999; 99DE-1041378.
 PR 31-AUG-1999; 99DE-1041379.
 PR 31-AUG-1999; 99DE-1041380.
 PR 31-AUG-1999; 99DE-1041394.
 PR 31-AUG-1999; 99DE-1041396.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042124.
 PR 03-SEP-1999; 99DE-1042129.
 PR 09-MAR-2000; 2000US-0187970.
 XX (BADI) BASF AG.
 XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauser G;
 XX WPI; 2001-137957/14.
 DR N-PSDB; AAF71937.
 XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
 PT pathway proteins, useful for producing fine chemicals in
 PT microorganisms, including organic acids, nonproteinogenic amino acids,
 PT and purine and pyrimidine bases -
 XX Claim 20; Page 711-714; 1737pp; English.
 PS AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
 CC MP nucleic acids are useful for the production of fine chemicals
 CC in microorganisms, including organic acids, nonproteinogenic amino
 CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
 CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
 CC compounds, vitamins, cofactors, polyketides and enzymes.
 XX Sequence 1066 AA;
 SQ Query Match 100.0%; Score 27; DB 22; Length 1066;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 |||||

Db 1061 DHAVK 1065

RESULT 15

AAB80134

ID AAB80134 standard; Protein; 1066 AA.

XX AAB80134;

AC AAB80134;

XX 30-APR-2001 (first entry)

DT DT

XX Corynebacterium glutamicum MP protein sequence SEQ ID NO:1002.

DE DE

XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;

KW fine chemical production; microorganism; organic acid; nucleoside;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;

KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;

KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.

XX

OS Corynebacterium glutamicum.

XX

XX WO200100843-A2.

XX 04-JAN-2001.

XX

PF 23-JUN-2000; 2000WO-IB00923.

XX

XX 25-JUN-1999; 99US-0141031.

PR 01-JUL-1999; 99DE-1030476.

PR 02-JUL-1999; 99US-0142101.

PR 08-JUL-1999; 99DE-1031415.

PR 08-JUL-1999; 99DE-1031418.

PR 08-JUL-1999; 99DE-1031419.

PR 08-JUL-1999; 99DE-1031420.

PR 08-JUL-1999; 99DE-1031424.

PR 08-JUL-1999; 99DE-1031428.

PR 08-JUL-1999; 99DE-1031434.

PR 08-JUL-1999; 99DE-1031435.

PR 08-JUL-1999; 99DE-1031443.

PR 08-JUL-1999; 99DE-1031453.

PR 08-JUL-1999; 99DE-1031457.

PR 08-JUL-1999; 99DE-1031465.

PR 08-JUL-1999; 99DE-1031478.

PR 08-JUL-1999; 99DE-1031510.

PR 08-JUL-1999; 99DE-1031541.

PR 08-JUL-1999; 99DE-1031573.

PR 08-JUL-1999; 99DE-1031632.

PR 08-JUL-1999; 99DE-1031634.

PR 08-JUL-1999; 99DE-1031636.

PR 08-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032126.

PR 09-JUL-1999; 99DE-1032130.

PR 09-JUL-1999; 99DE-1032186.

PR 09-JUL-1999; 99DE-1032206.

PR 09-JUL-1999; 99DE-1032227.

PR 09-JUL-1999; 99DE-1032228.

PR 09-JUL-1999; 99DE-1032229.

PR 09-JUL-1999; 99DE-1032230.

PR 14-JUL-1999; 99DE-1032922.

PR 14-JUL-1999; 99DE-1032926.

PR 14-JUL-1999; 99DE-1032928.

PR 14-JUL-1999; 99DE-1033004.

PR 14-JUL-1999; 99DE-1033005.

PR 14-JUL-1999; 99DE-1033006.

PR 12-AUG-1999; 99US-0148613.

PR 27-AUG-1999; 99DE-1040764.

PR 27-AUG-1999; 99DE-1040765.

PR 27-AUG-1999; 99DE-1040766.

PR 27-AUG-1999; 99DE-1040832.

PR 31-AUG-1999; 99DE-1041378.

PR 31-AUG-1999; 99DE-1041379.

PR 31-AUG-1999; 99DE-1041380.

PR 31-AUG-1999; 99DE-1041394.

PR 31-AUG-1999; 99DE-1041396.

PR 03-SEP-1999; 99DE-1042076.

PR 03-SEP-1999; 99DE-1042077.

PR 03-SEP-1999; 99DE-1042079.

PR 03-SEP-1999; 99DE-1042086.

PR 03-SEP-1999; 99DE-1042087.

PR 03-SEP-1999; 99DE-1042088.

PR 03-SEP-1999; 99DE-1042095.

PR 03-SEP-1999; 99DE-1042124.

PR 03-SEP-1999; 99DE-1042129.

PR 09-MAR-2000; 2000US-0187970.

XX

XX (BADI) BASF AG.

PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX WPI; 2001-137957/14..

XX N-PSDB; AAF72253.

DR

XX Nucleic acids from Corynebacterium glutamicum encoding metabolic

PT pathway proteins, useful for producing fine chemicals in

PT microorganisms, including organic acids, nonproteinogenic amino acids,

PT and purine and pyrimidine bases -

XX

PS Claim 20; Page 1540-1543; 1737pp; English.

XX

CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic

CC pathway (MP) proteins given in AAF79634 to AAF80211. The C. glutamicum

CC MP nucleic acids are useful for the production of fine chemicals

CC in microorganisms, including organic acids, nonproteinogenic amino

CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,

CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic

CC compounds, vitamins, cofactors, polyketides and enzymes.

XX

SQ Sequence 1066 AA;

Query Match 100.0%; Score 27; DB 22; Length 1066;

Best Local Similarity 100.0%; Pred. No. 66+02; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0;

OY 1 DHAVK 5

Db 1061 DHAVK 1065

RESULT 16

AAG91525

ID AAG91525 standard; Protein; 1113 AA.

XX

AC AAG91525;

XX

DT 26-SEP-2001 (first entry)

DE

XX C glutamicum protein fragment SEQ ID NO: 5279.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

XX

OS Corynebacterium glutamicum.

XX

PN EP1108790-A2.

XX

XX 20-JUN-2001.

XX

XX 18-DEC-2000; 2000EP-0127688.

XX

XX 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

XX

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI; 2001-376931/40.
 DR N-PSDB; AAF66744.
 DR
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX
 PS Claim 17; SEQ ID NO: 5279; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from corynebacterium bacterium, and identifying a homologue of a gene derived
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 XX Sequence 1113 AA;
 SQ
 Query Match 100.0%; Score 27; DB 22; Length 1113;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRAVK 5
 Db 1108 DRAVK 1112
 RESULT 17
 AAB79817
 ID AAB79817 standard; Protein; 1113 AA.
 XX
 AC AAB79817;
 DT
 DT 30-APR-2001 (first entry)
 XX
 DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:368.
 XX
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
 KW fine chemical production; microorganism; organic acid; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN WO200100843-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB00923.
 XX
 PR 25-JUN-1999; 99US-0141031.
 PR 01-JUL-1999; 99DE-1030476.
 PR 02-JUL-1999; 99US-0142101.
 PR 08-JUL-1999; 99DE-1031415.
 PR 08-JUL-1999; 99DE-1031418.
 PR 08-JUL-1999; 99DE-1031419.
 PR 08-JUL-1999; 99DE-1031420.
 PR 08-JUL-1999; 99DE-1031424.
 PR 08-JUL-1999; 99DE-1031428.
 PR 08-JUL-1999; 99DE-1031434.
 PR 08-JUL-1999; 99DE-1031435.

PR 08-JUL-1999; 99DE-1031443.
 PR 08-JUL-1999; 99DE-1031453.
 PR 08-JUL-1999; 99DE-1031457.
 PR 08-JUL-1999; 99DE-1031465.
 PR 08-JUL-1999; 99DE-1031478.
 PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031541.
 PR 08-JUL-1999; 99DE-1031573.
 PR 08-JUL-1999; 99DE-1031592.
 PR 08-JUL-1999; 99DE-1031632.
 PR 08-JUL-1999; 99DE-1031634.
 PR 08-JUL-1999; 99DE-1031636.
 PR 09-JUL-1999; 99DE-1032125.
 PR 09-JUL-1999; 99DE-1032126.
 PR 09-JUL-1999; 99DE-1032130.
 PR 09-JUL-1999; 99DE-1032186.
 PR 09-JUL-1999; 99DE-1032206.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032228.
 PR 09-JUL-1999; 99DE-1032230.
 PR 09-JUL-1999; 99DE-1032230.
 PR 14-JUL-1999; 99DE-1032922.
 PR 14-JUL-1999; 99DE-1032926.
 PR 14-JUL-1999; 99DE-1032928.
 PR 14-JUL-1999; 99DE-1033004.
 PR 14-JUL-1999; 99DE-1033005.
 PR 14-JUL-1999; 99DE-1033006.
 PR 12-AUG-1999; 99US-0148613.
 PR 27-AUG-1999; 99DE-1040764.
 PR 27-AUG-1999; 99DE-1040765.
 PR 27-AUG-1999; 99DE-1040766.
 PR 27-AUG-1999; 99DE-1040832.
 PR 31-AUG-1999; 99DE-1041378.
 PR 31-AUG-1999; 99DE-1041379.
 PR 31-AUG-1999; 99DE-1041380.
 PR 31-AUG-1999; 99DE-1041394.
 PR 31-AUG-1999; 99DE-1041396.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042077.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042124.
 PR 03-SEP-1999; 99DE-1042129.
 PR 09-MAR-2000; 2000US-0187970.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI; 2001-137957/14.
 DR N-PSDB; AAF71936.
 XX
 PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
 PT pathway proteins, useful for producing fine chemicals in
 PT microorganisms, including organic acids, nonproteinogenic amino acids,
 PT and purine and pyrimidine bases -
 XX
 PS Claim 20; Page 702-706; 1737pp; English.
 XX
 CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
 CC MP nucleic acids are useful for the production of fine chemicals
 CC in microorganisms, including organic acids, nonproteinogenic amino
 CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
 CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
 CC compounds, vitamins, cofactors, polyketides and enzymes.
 XX
 SQ Sequence 1113 AA;

Query Match

100.0%; Score 27; DB 22; Length 1113;

```
Best Local Similarity 100.0%; Pred. No. 6.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

QY 1 DHAUK 5
Db 1108 DHAUK 1112

RESULT 18
AAB80133
ID AAB80133 standard; Protein; 1113 AA.
XX
AC AAB80133;
XX
DT 30-APR-2001 (first entry)
DE Corynebacterium glutamicum MP protein sequence SRQ ID NO:1000.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00923.
XX
PR 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 08-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 08-JUL-1999; 99DE-1032125.
PR 08-JUL-1999; 99DE-1032126.
PR 08-JUL-1999; 99DE-1032130.
PR 08-JUL-1999; 99DE-1032186.
PR 08-JUL-1999; 99DE-1032206.
PR 08-JUL-1999; 99DE-1032227.
PR 08-JUL-1999; 99DE-1032228.
PR 08-JUL-1999; 99DE-1032229.
PR 08-JUL-1999; 99DE-1032230.
PR 08-JUL-1999; 99DE-1032232.
PR 08-JUL-1999; 99DE-1032292.
PR 08-JUL-1999; 99DE-1032926.
PR 08-JUL-1999; 99DE-1032928.
PR 08-JUL-1999; 99DE-1033004.
PR 08-JUL-1999; 99DE-1033005.
PR 08-JUL-1999; 99DE-1033006.
PR 08-JUL-1999; 99US-0148613.
PR 08-JUL-1999; 99DE-1040764.
PR 08-JUL-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BADI ) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX WPI; 2001-137957/14.
DR N-PSDB; AAF72252.
XX
PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 20; Page 1531-1534; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 1113 AA;
Query Match 100.0%; Score 27; DB 22; Length 1113;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAUK 5
Db 1108 DHAUK 1112

RESULT 19
AAU79297
ID AAU79297 standard; Protein; 1113 AA.
XX
AC AAU79297;
XX
DT 13-AUG-2002 (first entry)
XX
DE Polypeptide with carbamoylphosphate synthase activity #2.
XX
KW Corynebacterium; carbamoylphosphate synthase; enzyme; L-arginine.
XX
OS Brevibacterium lactofermentum.
XX
PN JP2002101881-A.
PD 09-APR-2002.
PF 31-JUL-2000; 2000JP-0231990.
PR 25-JUL-2000; 2000JP-0224082.
```

PA (AJIN) AJINOMOTO KK.
 XX WPI: 2002-447726/48.
 DR N-PSDB; ABK52437.
 XX
 PT Carbamoylphosphate synthase gene of a Corynebacterium and preparation
 of L-arginine
 XX
 PS Claim 1; Page 19-22; 24pp; Japanese.
 XX
 CC The invention relates to a Corynebacterium carbamoylphosphate synthase
 CC gene-related DNA fragment encoding polypeptides which have
 CC carbamoylphosphate synthase activity. The sequences are used in the
 CC preparation of L-arginine. This sequence represents a polypeptide with
 CC carbamoylphosphate synthase activity.
 XX
 SQ Sequence 1113 AA;

Query Match. 100.0%; Score 27; DB 23; Length 1113;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5
 |||||
 Db 1108 DHAVK 1112

RESULT 20
 ABG03370
 ID ABG03370 standard; Protein; 1245 AA.

XX AC ABG03370;
 XX
 DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #3361.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB; AAS67557.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX

PS Claim 20; SEQ ID NO 33729; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences, ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1245 AA;

Query Match. 100.0%; Score 27; DB 22; Length 1245;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5
 |||||
 Db 1100 DHAVK 1104

RESULT 21
 ABG14366
 ID ABG14366 standard; Protein; 1277 AA.

XX AC ABG14366;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #14357.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73,

DR N-PSDB; AAS78553.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX

PS Claim 20; SEQ ID NO 44725; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1277 AA;

Query Match 100.0%; Score 27; DB 22; Length 1277;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAUK 5
Db 1132 DHAUK 1136
|||||

RESULT 22

ABB64494
ID ABB64494 standard; Protein; 2248 AA.

XX AC ABB64494;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 20274.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX DT 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL08597.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX PS Disclosure; SEQ ID NO 20274; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2248 AA;

Query Match 100.0%; Score 27; DB 22; Length 2248;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAUK 5
Db 2 DHAUK 6
|||||

RESULT 23

AAW94614
ID AAW94614 standard; Protein; 167 AA.

XX AC AAW94614;

XX DT 29-APR-1999 (first entry)

XX DE Helicobacter pylori Pfr protein.

XX KW Helicobacter pylori; Pfr; intracellular cytosolic protein; vaccine;
XX KW immune response; infection; prophylaxis; diagnosis.

XX OS Helicobacter pylori.

XX PN WO9856815-A1.

XX PD 17-DEC-1998.

XX PF 08-JUN-1998; 98WO-SE01092;

XX PR 12-JUN-1997; 97SE-0002241.

XX PA (ASTR) ASTRA AB.

XX PI Berglindh T, Mellgard B;

XX DR WPI; 1999-080887/07.

XX DR N-PSDB; AAX16591.

XX PT New Helicobacter pylori Pfr polypeptide - useful for inducing a
PT protective immune response to H. pylori infection

XX PS Claim 2; Page 22; 34pp; English.

XX CC The present sequence represents a Helicobacter pylori Pfr protein, which
CC is an intracellular cytosolic protein. Pfr protein can be used in
CC inducing a protective immune response to H. pylori infection. The
CC protein forms a vaccine, which further comprises an adjuvant (a
CC pharmaceutically acceptable form of cholera toxin), and is used
CC therapeutically as a prophylactic vaccine in mammals (especially humans)
CC infected by H. pylori. It is additionally used to form a composition in
CC the treatment and diagnosis of H. pylori infection, and in eliciting a
CC protective immune response against such infection. Under stress,
CC H. pylori transforms from a bacillary form into a coccoidal form, which
CC is more resistant to antibiotics and other antibacterial agents. The
CC polypeptide elicits an immune response towards both the coccoidal and
CC bacillary forms of H. pylori. They also increase protective immune
CC mechanisms locally in the stomach, which is important because systemic
CC immunity only plays a limited role in protection against mucosal
CC infections.

XX SQ Sequence 167 AA;

Query Match 96.3%; Score 26; DB 20; Length 167;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAUK 5
Db 106 DHAUK 110
|||||

RESULT 24

AAW89864
ID AAW89864 standard; Protein; 167 AA.

XX AC AAW89864;
XX DT 18-FEB-1999 (first entry)
XX DE Antigen 1 from cluster 11a.
XX KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
XX KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
XX OS Helicobacter pylori.

XX PN WO9849314-A2.

XX PD 05-NOV-1998.

XX PF 27-APR-1998; 98WO-US08487.

XX PR 14-OCT-1997; 97US-0061958.

XX PR 25-APR-1997; 97US-0045107.

XX PA (GENE-) GENELABS TECHNOLOGIES INC.

XX PI Chow TP, Fry KE, Lim MY, McAtee CP;

XX DR WPI; 1999-009433/01.

XX PT New Helicobacter pylori antigens and related nucleic acid sequences.
XX PT - useful in serological diagnosis and protective vaccines, providing
XX PT long-lasting immune response
XX PS Claim 1; Page 218; 402pp; English.

XX CC The present sequence represents a Helicobacter pylori antigenic protein
XX CC that is characterised by immunoreactivity with H. pylori-positive
XX CC antisera. The proteins are highly immunogenic and induce a long-lasting
XX CC immune response that persists even after antimicrobial treatment. In
XX CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
XX CC highly sensitive and specific. The specification also describes 69
XX CC previously unrecognised immunogenic cluster families. H. pylori antigens
XX CC are used to detect H. pylori-specific antibodies, for diagnosing
XX CC infection or to confirm eradication of infection, and in vaccines to
XX CC protect against H. pylori infection and related diseases (gastritis,
XX CC peptic ulcer, gastric adenocarcinoma/lymphoma).

SQ Sequence 167 AA;

Query Match 96.3%; Score 26; DB 20; Length 167;

Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5

Db 106 DHAIK 110

RESULT 25

AAW89864
ID AAW89864 standard; Protein; 167 AA.

XX AC AAW89864;
XX DT 04-DEC-2000 (first entry)
XX DE Amino acid sequence of a 19 kDa protein of Helicobacter pylori.
XX KW 19 kDa protein; vaccine; immune response; bacterial infection;
XX KW Helicobacter infection.
XX OS Helicobacter pylori.

XX FH Key Location/Qualifiers
XX FT Misc-difference 114 /note= "His encoded by AT"

XX PN WO2000046242-A2.

XX PD 10-AUG-2000.

XX PF 03-FEB-2000; 2000WO-US02938.

XX PR 04-FEB-1999; 99US-0118631.

XX PA (AMCY) AMERICAN CYANAMID CO.

XX PI Fiske M, Zhu D, Fulginti JP, Schmidt SG;

XX DR WPI; 2000-514947/46.

XX DR N-PSDB; AAA63814.

XX PT Novel vaccine composition for treating Helicobacter infection
XX PT comprising an isolated and purified polypeptide of Helicobacter pylori
XX PT with a molecular weight of 19 kDa
XX PS Disclosure; Page 36-37; 39pp; English.

XX CC The present sequence represents a 19 kDa protein of Helicobacter pylori.
XX CC The protein is used in vaccine compositions. These vaccine compositions
XX CC are useful for inducing an immune response in a mammal. The 19 kDa
XX CC polypeptide is useful for diagnostic and therapeutic purposes such
XX CC as detecting the presence of Helicobacter pylori and other
XX CC Helicobacter species in a sample, to screen compounds having the
XX CC ability to interfere the bacterial life cycle or inhibit bacterial
XX CC infection, and for treating and preventing Helicobacter infections.

SQ Sequence 167 AA;

Query Match 96.3%; Score 26; DB 21; Length 167;

Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5

Db 106 DHAIK 110

Search completed: June 20, 2003, 20:51:55

Job time : 34.9318 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 10.5682 Seconds
(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-14

Perfect score: 27

Sequence: 1 DRAVK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/aaa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/aaa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/aaa/5A_COMB.pep:*
- 4: /cgn2_6/prodata/1/aaa/5B_COMB.pep:*
- 5: /cgn2_6/prodata/1/aaa/PTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	5	4	US-08-996-679-55
2	27	100.0	5	4	/cgn2_6/prodata/1/aaa/5A_COMB.pep:*
3	27	100.0	5	4	/cgn2_6/prodata/1/aaa/5B_COMB.pep:*
4	27	100.0	5	4	/cgn2_6/prodata/1/aaa/5A_COMB.pep:*
5	27	100.0	5	4	/cgn2_6/prodata/1/aaa/5B_COMB.pep:*
6	27	100.0	5	4	US-09-458-870-14
7	27	100.0	1113	4	US-09-629-616-3
8	24	88.9	193	1	US-08-260-202A-19
9	24	88.9	193	1	US-08-017-114-19
10	24	88.9	193	3	US-08-505-307-19
11	24	88.9	193	4	US-09-082-920-4
12	24	88.9	193	4	US-09-609-151A-19
13	24	88.9	193	5	PCT-US94-02034-19
14	24	88.9	383	1	US-08-486-037B-2
15	24	88.9	564	1	US-07-872-644-53
16	24	88.9	564	1	US-08-297-494-53
17	24	88.9	564	1	US-08-297-510-53
18	24	88.9	564	1	US-08-479-526-53
19	24	88.9	564	1	US-08-455-526-53
20	24	88.9	564	1	US-08-455-525-53
21	24	88.9	564	3	US-09-139-491-53
22	24	88.9	564	5	PCT-US92-03222-53
23	24	88.9	593	6	5523211-2
24	24	88.9	615	3	US-09-042-426-10
25	24	88.9	615	4	US-09-291-238-10
26	24	88.9	615	4	US-09-330-760-10
27	24	88.9	615	4	US-09-328-473-10

ALIGNMENTS

RESULT 1

US-08-996-679-55

; Sequence 55, Application US/08996679

; Patent No. 6169071

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING

; TITLE OF INVENTION: CELL ADHESION

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED AND BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

28	24	88.9	615	4	US-09-330-737-10	Sequence 10, Appl
29	24	88.9	615	4	US-09-329-169-10	Sequence 10, Appl
30	24	88.9	615	4	US-09-330-714A-10	Sequence 10, Appl
31	24	88.9	615	4	US-09-328-826-10	Sequence 10, Appl
32	24	88.9	634	1	US-07-872-644-51	Sequence 51, Appl
33	24	88.9	634	1	US-08-297-494-51	Sequence 51, Appl
34	24	88.9	634	1	US-08-297-510-51	Sequence 51, Appl
35	24	88.9	634	1	US-08-479-532-51	Sequence 51, Appl
36	24	88.9	634	1	US-08-455-526-51	Sequence 51, Appl
37	24	88.9	634	1	US-08-455-525-51	Sequence 51, Appl
38	24	88.9	634	3	PCT-US92-03222-51	Sequence 51, Appl
39	24	88.9	695	5	PCT-US92-03222-51	Sequence 51, Appl
40	24	88.9	969	1	US-07-671-817A-5	Sequence 6, Appl
41	24	88.9	969	1	US-07-671-817A-6	Sequence 6, Appl
42	24	88.9	1009	2	US-08-680-326-31	Sequence 31, Appl
43	24	88.9	1074	4	US-09-004-838-11	Sequence 11, Appl
44	24	88.9	1155	1	US-08-349-867-33	Sequence 33, Appl
45	24	88.9	1155	1	US-08-446-486-2	Sequence 2, Appl
46	24	88.9	1155	1	US-07-951-715A-9	Sequence 9, Appl
47	24	88.9	1155	1	US-08-463-308-2	Sequence 2, Appl
48	24	88.9	1155	2	US-08-598-308A-33	Sequence 33, Appl
49	24	88.9	1155	2	US-08-459-448A-9	Sequence 9, Appl
50	24	88.9	1155	3	US-08-459-595A-9	Sequence 9, Appl
51	24	88.9	1155	3	US-09-021-203-2	Sequence 2, Appl
52	24	88.9	1155	3	US-08-459-504B-9	Sequence 9, Appl
53	24	88.9	1155	3	US-08-459-444-9	Sequence 9, Appl
54	24	88.9	1155	3	US-09-053-549-4	Sequence 4, Appl
55	24	88.9	1155	3	US-09-053-549-6	Sequence 6, Appl
56	24	88.9	1155	4	US-09-547-422-9	Sequence 9, Appl
57	24	88.9	1155	6	5254799-4	Patent No. 5254799
58	24	88.9	1165	1	US-08-239-476-33	Sequence 33, Appl
59	24	88.9	1165	2	US-08-639-923A-33	Sequence 33, Appl
60	24	88.9	1165	5	PCT-US95-05431-33	Sequence 33, Appl
61	24	88.9	1177	3	US-08-754-490-14	Sequence 14, Appl
62	24	88.9	1177	3	US-08-754-490-26	Sequence 26, Appl
63	24	88.9	1177	3	US-08-923-505A-14	Sequence 14, Appl
64	24	88.9	1177	3	US-08-923-505A-26	Sequence 26, Appl
65	24	88.9	1177	4	US-09-260-952A-14	Sequence 14, Appl
66	24	88.9	1177	4	US-09-260-952A-26	Sequence 26, Appl
67	24	88.9	1177	4	US-09-253-341-14	Sequence 14, Appl
68	24	88.9	1177	4	US-09-253-341-26	Sequence 26, Appl
69	24	88.9	1177	4	US-09-253-331A-14	Sequence 14, Appl
70	24	88.9	1177	4	US-09-253-331A-26	Sequence 26, Appl
71	24	88.9	1177	4	US-09-261-040-14	Sequence 14, Appl
72	24	88.9	1177	4	US-09-261-040-26	Sequence 26, Appl
73	24	88.9	1181	1	US-07-951-715A-11	Sequence 11, Appl
74	24	88.9	1181	1	US-07-951-715A-13	Sequence 13, Appl
75	24	88.9	1181	1	US-07-951-715A-15	Sequence 15, Appl

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-55

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 1 DHAVK 5

RESULT 2

US-09-115-395-17
Sequence 17, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115,395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996,679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893,534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 17
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis

US-09-115-395-17

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 1 DHAVK 5

RESULT 3

US-09-250-059-14
Sequence 14, Application US/09250059
Patent No. 633307
GENERAL INFORMATION:

APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250,059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-250-059-14

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 1 DHAVK 5

RESULT 4

US-09-248-074-14
Sequence 14, Application US/09248074
Patent No. 6346512
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
FILE REFERENCE: 100086.401C5
CURRENT APPLICATION NUMBER: US/09/248,074
CURRENT FILING DATE: 1999-02-10
NUMBER OF SEQ ID NOS: 81
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-248-074-14

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 1 DHAVK 5

RESULT 5

US-09-357-717-13
Sequence 13, Application US/09357717
Patent No. 6417325

GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farooqui, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with cadherin cell adhesion recognition
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
; US-09-357-717-13

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
| | | | |
DB 1 DHAVK 5

RESULT 6
US-09-458-870-14
; Sequence 14, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farooqui, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with cadherin cell adhesion recognition
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
; US-09-458-870-14

Query Match 100.0%; Score 27; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
| | | | |
DB 1 DHAVK 5

RESULT 7
US-09-629-616-3
; Sequence 3, Application US/09629616
; Patent No. 6255086
; GENERAL INFORMATION:
; APPLICANT: KAWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM
; TITLE OF INVENTION: BACTERIA AND METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: OP945CIP
; CURRENT APPLICATION NUMBER: US/09/629,616
; CURRENT FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
; US-09-629-616-3

Query Match 100.0%; Score 27; DB 4; Length 1113;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
| | | | |
DB 1108 DHAVK 1112

RESULT 8
US-08-260-202A-19
; Sequence 19, Application US/08260202A
; Patent No. 5573910
; GENERAL INFORMATION:
; APPLICANT: Deretic, Vojo
; APPLICANT: Martin, Daniel W.
; TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOIDY IN
; TITLE OF INVENTION: PSEUDOMONAS AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/260,202A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,202
; FILING DATE: 15-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/017,114
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hodgins, Daniel S.
; REGISTRATION NUMBER: 31,026
; REFERENCE/DOCKET NUMBER: UTSK:221\HOD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-260-202A-19

Query Match 88.9%; Score 24; DB 1; Length 193;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 110 DHALK 114

RESULT 9

US-08-017-114-19
; Sequence 19, Application US/08017114
; Patent No. 5591838
; GENERAL INFORMATION:
; APPLICANT: Deretic, Vojo
; TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOIDY IN
; TITLE OF INVENTION: PSEUDOMONAS AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENTS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19930212
; APPLICATION NUMBER: US/08/017,114
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hodgins, Daniel S.
; REGISTRATION NUMBER: 31,026
; REFERENCE/DOCKET NUMBER: UTSK:205/HOD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/320-7200
; TELEFAX: 512/474-7577
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-017-114-19

Query Match 88.9%; Score 24; DB 1; Length 193;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 110 DHALK 114

RESULT 10

US-08-505-307-19
; Sequence 19, Application US/08505307
; Patent No. 6083691
; GENERAL INFORMATION:

APPLICANT: Deretic, Vojo
APPLICANT: Martin, Daniel W.
TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOIDY IN
TITLE OF INVENTION: PSEUDOMONAS AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,307
FILING DATE: 24-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/017,114
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hodgins, Daniel S.
REGISTRATION NUMBER: 31,026
REFERENCE/DOCKET NUMBER: UTSK:231
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-505-307-19

Query Match 88.9%; Score 24; DB 3; Length 193;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 110 DHALK 114

RESULT 11

US-09-082-920-4
; Sequence 4, Application US/09082920C
; Patent No. 6355469
; GENERAL INFORMATION:
; APPLICANT: Lam, Kelvin T.
; TITLE OF INVENTION: Nucleic Acid Encoding M. Tuberculosis ALGU Protein
; FILE REFERENCE: 0342/1C587USIAPPEND
; CURRENT APPLICATION NUMBER: US/09/082,920C
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 60/035,391
; EARLIER FILING DATE: 1997-01-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-082-920-4

Query Match 88.9%; Score 24; DB 4; Length 193;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5

Db 110 DHALK 114
|||:|

RESULT 12

US-09-609-151A-19
; Sequence 19, Application US/09609151A
; Patent No. 6426187
; GENERAL INFORMATION:
; APPLICANT: DERETIC, VOJO
; APPLICANT: MARTIN, DANIEL W.
; TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOIDY IN PSEUDOMONAS
; FILE REFERENCE: UTSK:231USC1
; CURRENT APPLICATION NUMBER: US/09/609,151A
; CURRENT FILING DATE: 2000-06-30
; PRIOR FILING DATE: 1995-11-24
; PRIOR APPLICATION NUMBER: 08/505,307
; PRIOR FILING DATE: 1993-02-12
; PRIOR APPLICATION NUMBER: 08/017,114
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-609-151A-19

Query Match 88.9%; Score 24; DB 4; Length 193;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||:|
Db 110 DHALK 114

RESULT 13

PCT-US94-02034-19
; Sequence 19, Application PC/TUS9402034
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: STREET: 201 West 7th Street
; APPLICANT: CITY: Austin
; APPLICANT: STATE: Texas
; APPLICANT: COUNTRY: United States of America
; APPLICANT: POSTAL CODE: 78701
; APPLICANT: TELEPHONE NO: (512)499-4462
; APPLICANT: TELEFAX: (512)499-4523
; TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOIDY IN PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENTS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/02034
; FILING DATE: UNKNOWN
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/017,114
; FILING DATE: 12 February 1993 (12.02.93)
; ATTORNEY/AGENT INFORMATION:
; NAME: HODGINS, DANIEL S.
; REGISTRATION NUMBER: 31,026

REFERENCE/DOCKET NUMBER: UTFK231P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/320-7200
TELEFAX: 713/789-2679
TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US94-02034-19

Query Match 88.9%; Score 24; DB 5; Length 193;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||:|
Db 110 DHALK 114

RESULT 14

US-08-486-037B-2
; Sequence 2, Application US/08486037B
; Patent No. 5607672
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jeffrey D.
; TITLE OF INVENTION: REPLACEMENT THERAPY FOR DENTAL CARIES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,037B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07397/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 383 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-037B-2

Query Match 88.9%; Score 24; DB 1; Length 383;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||:|
Db 355 DHALK 359

RESULT 15

US-07-872-644-53

; Sequence 53, Application US/07872644
; Patent No. 5389527
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,644
; FILING DATE: 19920420
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5389527and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-872-644-53

Query Match 88.9%; Score 24; DB 1; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 37 DHALK 41

RESULT 16
US-08-297-494-53
; Sequence 53, Application US/08297494
; Patent No. 5580771
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street

; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,494
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5580771and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-297-494-53

Query Match 88.9%; Score 24; DB 1; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 37 DHALK 41

RESULT 17
US-08-297-510-53
; Sequence 53, Application US/08297510
; Patent No. 5602019
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,510
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356

;; FILING DATE: 04-APR-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5602019and, Greta E.
;; REGISTRATION NUMBER: 35,302
;; REFERENCE/DOCKET NUMBER: 278666/30822
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 346-5750
;; TELEFAX: (312) 984-9740
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 53:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 564 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-297-510-53

Query Match 88.9%; Score 24; DB 1; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:|
DB 37 DHALK 41

RESULT 18
US-08-479-532-53
; Sequence 53, Application US/08479532
; Patent No. 5776752
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/479,532
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/297,494
; FILING DATE:
; APPLICATION NUMBER: 08/688,356
; FILING DATE:
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5776752and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 278666/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid

;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-479-532-53
Query Match 88.9%; Score 24; DB 1; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
|||:|
DB 37 DHALK 41
RESULT 19
US-08-455-526-53
; Sequence 53, Application US/08455526
; Patent No. 5789553
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/455,526
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/297,494
; FILING DATE: 29-AUG-1994
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5789553and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 278666/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-526-53

Query Match 88.9%; Score 24; DB 1; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:|
DB 37 DHALK 41

RESULT 20
US-08-455-525-53
; Sequence 53, Application US/08455525
; Patent No. 5800987
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,525
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/297,494
; FILING DATE:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5800987and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-525-53

Query Match 88.9%; Score 24; DB 1; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 37 DHALK 41

RESULT 21
US-09-139-491-53
; Sequence 53, Application US/09139491
; Patent No. 6015677
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/139,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,525
; FILING DATE: 31-MAY-1995
; APPLICATION NUMBER: 08/297,494
; FILING DATE:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6015677and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-139-491-53

Query Match 88.9%; Score 24; DB 3; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 37 DHALK 41

RESULT 22
PCT-US92-03222-53
; Sequence 53, Application PC/TUS9203222
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03222
FILING DATE: 19920420
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,356
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-03222-53

Query Match 88.9%; Score 24; DB 5; Length 564;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 37 DHALK 41

RESULT 23
5523211-2
; Patent No. 5523211
; APPLICANT: PUSZTAI-CAREY, MARIANNE; CAREY, PAUL R.; LESSARD,
; TIMOTHY; YAGUCHI, MAKOTO
; TITLE OF INVENTION: IDENTIFICATION, QUANTITATION AND
; PURIFICATION OF INSECTICIDAL PROTEINS FROM BACILLUS THURINGIENSIS
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/277,076
; FILING DATE: 19-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 102,491
; FILING DATE: 05-AUG-1993
; APPLICATION NUMBER: 836,967
; FILING DATE: 19-FEB-1992
; APPLICATION NUMBER: 493,453
; FILING DATE: 14-MAR-1990
; SEQ ID NO: 2:
; LENGTH: 593
5523211-2

Query Match 88.9%; Score 24; DB 6; Length 593;
Best Local Similarity 80.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 176 DHAVR 180

RESULT 24
US-09-042-426-10
; Sequence 10, Application US/09042426
; Patent No. 6114608
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6114608artis Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,426
FILING DATE: March 13, 1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Hoxie, Thomas
REGISTRATION NUMBER: 32,993
REFERENCE/DOCKET NUMBER: 135/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8614
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 615 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
IMMEDIATE SOURCE:
CLONE: Bt protein
US-09-042-426-10

Query Match 88.9%; Score 24; DB 3; Length 615;
Best Local Similarity 80.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 205 DHAVR 209

RESULT 25
US-09-291-238-10
; Sequence 10, Application US/09291238
; Patent No. 6222104
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6222104artis Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/291,238
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/042,426
FILING DATE: March 13, 1998

ATTORNEY/AGENT INFORMATION:

NAME: Hoxie, Thomas
REGISTRATION NUMBER: 32,993
REFERENCE/DOCKET NUMBER: 135/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8614
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 615 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
IMMEDIATE SOURCE:
CLONE: Bt protein
US-09-291-238-10

Query Match 88.9%; Score 24; DB 4; Length 615;

Best Local Similarity 80.0%; Pred. No. 6.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5

Db 205 DHAVR 209

Search completed: June 20, 2003, 21:00:30
Job time : 11.5682 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:52:19 ; Search time 12.8409 Seconds
(without alignments)
42.134 Million cell updates/sec

Title: US-10-105-008-14
Perfect score: 27
Sequence: 1 DHAVK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Published Applications AA:**
1: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05 NEW PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US03 PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10 PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	5	9	US-09-769-145-14
2	27	100.0	5	9	US-10-058-821-13
3	27	100.0	5	10	US-09-234-395-310
4	27	100.0	5	10	US-09-305-928-310
5	27	100.0	459	9	US-10-007-527A-21
6	27	100.0	459	9	US-10-007-452-21
7	27	100.0	508	1	US-08-781-986A-5241
8	27	100.0	510	9	US-09-738-626-5069
9	27	100.0	1113	9	US-09-738-626-5279
10	27	100.0	1113	9	US-10-284-138-3
11	27	100.0	1113	9	US-10-284-334-3
12	27	100.0	1113	10	US-09-836-470B-3
13	27	100.0	1324	10	US-09-801-368-222
14	26	96.3	515	10	US-09-815-242-5811
15	26	96.3	525	10	US-09-815-242-12926
16	26	96.3	2492	10	US-09-991-258-3
17	24	88.9	43	10	US-09-864-761-46135
18	24	88.9	83	10	US-09-734-569-84
19	24	88.9	88	9	US-09-895-913A-82

ALIGNMENTS

RESULT 1:
US-09-769-145-14
; Sequence 14, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Amar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoaemeng

20	24	88.9	143	10	US-09-739-907-56	Sequence 56, Appl
21	24	88.9	151	10	US-09-739-907-102	Sequence 102, Appl
22	24	88.9	176	12	US-10-062-254-96	Sequence 96, Appl
23	24	88.9	222	12	US-10-062-254-88	Sequence 88, Appl
24	24	88.9	250	12	US-10-062-254-90	Sequence 90, Appl
25	24	88.9	295	12	US-10-062-254-100	Sequence 100, Appl
26	24	88.9	308	12	US-10-062-254-104	Sequence 104, Appl
27	24	88.9	353	12	US-10-062-254-98	Sequence 98, Appl
28	24	88.9	359	10	US-09-881-752A-248	Sequence 248, Appl
29	24	88.9	360	10	US-09-815-242-11598	Sequence 11598, A
30	24	88.9	453	10	US-09-758-623-2	Sequence 2, Appli
31	24	88.9	480	9	US-09-738-626-4560	Sequence 4560, Ap
32	24	88.9	564	10	US-09-883-825-53	Sequence 53, Appl
33	24	88.9	565	12	US-10-062-254-92	Sequence 92, Appl
34	24	88.9	567	9	US-09-344-882-30	Sequence 30, Appl
35	24	88.9	567	9	US-10-293-865-30	Sequence 30, Appl
36	24	88.9	570	9	US-09-344-882-28	Sequence 28, Appl
37	24	88.9	570	9	US-10-293-865-28	Sequence 28, Appl
38	24	88.9	570	12	US-10-062-254-102	Sequence 102, Appl
39	24	88.9	570	12	US-10-062-254-367	Sequence 367, Appl
40	24	88.9	634	10	US-09-883-825-51	Sequence 51, Appl
41	24	88.9	640	9	US-10-269-848-2	Sequence 2, Appli
42	24	88.9	824	9	US-09-909-567B-53	Sequence 53, Appl
43	24	88.9	912	9	US-09-291-417-26	Sequence 26, Appl
44	24	88.9	950	10	US-09-815-242-11961	Sequence 11961, A
45	24	88.9	968	9	US-09-291-417-107	Sequence 107, Appl
46	24	88.9	1139	9	US-09-971-490-5	Sequence 5, Appli
47	24	88.9	1155	9	US-09-988-462-9	Sequence 9, Appli
48	24	88.9	1155	10	US-09-756-643-2	Sequence 2, Appli
49	24	88.9	1177	10	US-09-873-873-26	Sequence 26, Appl
50	24	88.9	1181	9	US-09-988-462-11	Sequence 11, Appl
51	24	88.9	1181	9	US-09-988-462-13	Sequence 13, Appl
52	24	88.9	1181	9	US-09-988-462-15	Sequence 15, Appl
53	24	88.9	1181	9	US-09-988-462-17	Sequence 17, Appl
54	24	88.9	1181	9	US-09-988-462-28	Sequence 28, Appl
55	24	88.9	2431	10	US-09-901-106-2	Sequence 2, Appli
56	23	85.2	5	9	US-09-769-145-15	Sequence 15, Appl
57	23	85.2	15	9	US-09-903-327A-19	Sequence 19, Appl
58	23	85.2	32	9	US-10-174-410-205	Sequence 205, App
59	23	85.2	69	10	US-09-764-877-1739	Sequence 1739, Ap
60	23	85.2	85	10	US-09-987-108-19	Sequence 19, Appl
61	23	85.2	165	9	US-09-738-626-5314	Sequence 5314, Ap
62	23	85.2	241	9	US-09-978-248-11	Sequence 11, Appl
63	23	85.2	245	9	US-09-372-348-13	Sequence 13, Appl
64	23	85.2	337	9	US-09-738-626-3693	Sequence 3693, Ap
65	23	85.2	358	10	US-09-934-901-8	Sequence 8, Appli
66	23	85.2	358	10	US-09-934-968-18	Sequence 18, Appl
67	23	85.2	359	10	US-09-815-242-11040	Sequence 11040, A
68	23	85.2	370	9	US-09-738-626-5196	Sequence 5196, Ap
69	23	85.2	375	10	US-09-731-872-388	Sequence 388, App
70	23	85.2	375	12	US-10-003-806-4	Sequence 4, Appli
71	23	85.2	380	9	US-09-738-626-6372	Sequence 6372, Ap
72	23	85.2	438	9	US-10-012-055-5	Sequence 5, Appli
73	23	85.2	520	10	US-09-992-901-2	Sequence 56, Appl
74	23	85.2	525	9	US-10-028-072-56	Sequence 56, Appl
75	23	85.2	525	9	US-10-121-049-56	Sequence 56, Appl

APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 10086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-14

Query Match 100.0%; Score 27; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 1 DHAVK 5

RESULT 2
US-10-058-821-13
Sequence 13, Application US/10058821
Publication No. US20030087811A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
FILE REFERENCE: 10086.401C12
CURRENT APPLICATION NUMBER: US/10/058,821
CURRENT FILING DATE: 2002-01-29
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-10-058-821-13

Query Match 100.0%; Score 27; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 1 DHAVK 5

RESULT 3
US-09-234-395-310
Sequence 310, Application US/09234395
Patent No. US20020123044A1

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 10086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 310
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-234-395-310

Query Match 100.0%; Score 27; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 1 DHAVK 5

RESULT 4
US-09-305-928-310
Sequence 310, Application US/09305928
Patent No. US20020146687A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 10086.407C4
CURRENT APPLICATION NUMBER: US/09/305,928
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 310
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-305-928-310

Query Match 100.0%; Score 27; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 1 DHAVK 5

RESULT 5
US-10-007-527A-21
Sequence 21, Application US/10007527A
Publication No. US20030044807A1
GENERAL INFORMATION:
APPLICANT: Tomb, Jean-Francois
APPLICANT: Bramucci, Michael G.

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 10086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 310
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-234-395-310

Query Match 100.0%; Score 27; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 1 DHAVK 5

RESULT 4
US-09-305-928-310
Sequence 310, Application US/09305928
Patent No. US20020146687A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 10086.407C4
CURRENT APPLICATION NUMBER: US/09/305,928
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 310
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-305-928-310

Query Match 100.0%; Score 27; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 1 DHAVK 5

RESULT 5
US-10-007-527A-21
Sequence 21, Application US/10007527A
Publication No. US20030044807A1
GENERAL INFORMATION:
APPLICANT: Tomb, Jean-Francois
APPLICANT: Bramucci, Michael G.

```
/ APPLICANT: Cheng, Qiong
/ APPLICANT: Kostichka, Kristy N.
/ TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
/ FILE REFERENCE: CL1709 US NA
/ CURRENT APPLICATION NUMBER: US/10/007,527A
/ PRIOR FILING DATE: 2001-12-05
/ PRIOR APPLICATION NUMBER: 60/254,868
/ PRIOR FILING DATE: 2000-12-12
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 21
/ LENGTH: 459
/ TYPE: PRT
/ ORGANISM: Arcanobacterium pyogenes
US-10-007-527A-21

Query Match      100.0%; Score 27; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DHAVK 5
Db      169 DHAVK 173

RESULT 6
US-10-007-452-21
/ Sequence 21, Application US/10007452
/ Publication No. US20030093701A1
/ GENERAL INFORMATION:
/ APPLICANT: Tomb, Jean-Francois
/ APPLICANT: Bramucci, Michael G.
/ APPLICANT: Cheng, Qiong
/ APPLICANT: Kostichka, Kristy N.
/ TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
/ FILE REFERENCE: CL1709 US NA
/ CURRENT APPLICATION NUMBER: US/10/007,452
/ CURRENT FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: 60/254,868
/ PRIOR FILING DATE: 2000-12-12
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 21
/ LENGTH: 459
/ TYPE: PRT
/ ORGANISM: Arcanobacterium pyogenes
US-10-007-452-21

Query Match      100.0%; Score 27; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DHAVK 5
Db      169 DHAVK 173

RESULT 7
US-08-781-986A-5241
/ Sequence 5241, Application US/08781986A
/ Publication No. US20030054436A1
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 5255
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Human Genome Sciences, Inc..
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION NUMBER: US/08/781,986A
/ APPLICATION DATE:
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 5241:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 508 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-781-986A-5241

Query Match      100.0%; Score 27; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DHAVK 5
Db      66 DHAVK 70

RESULT 8
US-09-738-626-5069
/ Sequence 5069, Application US/09738626
/ Publication No. US20020197605A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKAGAWA, SATOSHI
/ APPLICANT: MIZOGUCHI, HIROSHI
/ APPLICANT: ANDO, SEIKO
/ APPLICANT: HAYASHI, MIKIRO
/ APPLICANT: OCHIAI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAKO
/ APPLICANT: SENOH, AKIHIRO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
/ SOFTWARE: PatentIn ver. 3.0
/ SEQ ID NO 5069
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Corynebacterium glutamicum
US-09-738-626-5069

Query Match      100.0%; Score 27; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 DHAVK 5
|||||
Db 418 DHAVK 422

RESULT 9

US-09-738-626-5279
; Sequence 5279, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738.626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5279
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5279

Query Match 100.0%; Score 27; DB 9; Length 1113;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||||
Db 1108 DHAVK 1112

RESULT 10

US-10-284-138-3
; Sequence 3, Application US/10284138
; Publication No. US20030082774A1
; GENERAL INFORMATION:
; APPLICANT: KUWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM
; TITLE OF INVENTION: BACTERIA AND METHOD FOR
; TITLE OF INVENTION: PRODUCING L-ARGININE
; FILE REFERENCE: 0010-1082-0
; CURRENT APPLICATION NUMBER: US/10/284.138
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US/09/494.359
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: JP 11-24149
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRT

; ORGANISM: Brevibacterium lactofermentum
US-10-284-138-3

Query Match 100.0%; Score 27; DB 9; Length 1113;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||||
Db 1108 DHAVK 1112

RESULT 11

US-10-284-334-3
; Sequence 3, Application US/10284334
; Publication No. US20030082775A1
; GENERAL INFORMATION:
; APPLICANT: KUWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM
; TITLE OF INVENTION: BACTERIA AND METHOD FOR
; TITLE OF INVENTION: PRODUCING L-ARGININE
; FILE REFERENCE: 0010-1082-0
; CURRENT APPLICATION NUMBER: US/10/284.334
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US/09/494.359
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: JP 11-24149
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-10-284-334-3

Query Match 100.0%; Score 27; DB 9; Length 1113;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||||
Db 1108 DHAVK 1112

RESULT 12

US-09-836-470B-3
; Sequence 3, Application US/09836470B
; Patent No. US20020090702A1
; GENERAL INFORMATION:
; APPLICANT: KUWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM BACTERIA
; TITLE OF INVENTION: PRODUCING L-ARGININE
; FILE REFERENCE: 205801USODIV
; CURRENT APPLICATION NUMBER: US/09/836.470B
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 09/629.616
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 09/494.359
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: JP 11-24149
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-09-836-470B-3

Query Match 100.0%; Score 27; DB 10; Length 1113;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5
Db 1108 DHAVK 1112

RESULT 13
US-09-801-368-222
; Sequence 222, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 222
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-222

Query Match 100.0%; Score 27; DB 10; Length 1224;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5
Db 68 DHAVK 72

RESULT 14
US-09-815-242-5811
; Sequence 5811, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5811
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5811

Query Match 96.3%; Score 26; DB 10; Length 515;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5
Db 151 DHAVK 155

RESULT 15
US-09-815-242-12926
; Sequence 12926, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12926
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Staphylococcus aureus

US-09-815-242-12926

Query Match 96.3%; Score 26; DB 10; Length 525;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||:
Db 161 DHA1K 165

RESULT 16

US-09-991-258-3
; Sequence 3, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swenstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 0113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991.258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
; OTHER INFORMATION: synthetic construct
US-09-991-258-3

Query Match 96.3%; Score 26; DB 10; Length 2492;
Best Local Similarity 80.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||:
Db 1230 DHA1K 1234

RESULT 17

US-09-864-761-46135
; Sequence 46135, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46135
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002036.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EST HUMAN HIT: AUI24141.1, EVALUE 6.00e-10
; OTHER INFORMATION: SWISSPROT HIT: P57080, EVALUE 2.00e-11
US-09-864-761-46135

Query Match 88.9%; Score 24; DB 10; Length 43;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||:
Db 35 DHSVK 39

RESULT 18

US-09-734-569-84
; Sequence 84, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins inv
; TITLE OF INVENTION: in the synthesis of carbohydrates
; FILE REFERENCE: BASF-NAE-1332-99-US

; CURRENT APPLICATION NUMBER: US/09/734,569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patentin Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 84
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-84

Query Match 88.9%; Score 24; DB 10; Length 83;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 43 DHALK 47

RESULT 19
US-09-895-913A-82
; Sequence 82, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-82

Query Match 88.9%; Score 24; DB 9; Length 88;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 46 DHALK 50

RESULT 20
US-09-739-907-56
; Sequence 56, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: PZ022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704

; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 56
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals stop translation
US-09-739-907-56

Query Match 88.9%; Score 24; DB 10; Length 143;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 114 DHSVK 118

RESULT 21
US-09-739-907-102
; Sequence 102, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: PZ022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 102
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-907-102

Query Match 88.9%; Score 24; DB 10; Length 151;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 123 DHSVK 127

RESULT 22
US-10-062-254-96
; Sequence 96, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen

APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/062.254
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 09/630,346
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/146511
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/156006
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/156899
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 60/157287
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/169767
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171054
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/172958
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/171515
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 60/173535
NUMBER OF SEQ ID NOS: 375
SOFTWARE: Microsoft Office 97
SEQ ID NO 96
LENGTH: 176
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: UNSURE
LOCATION: (92)
US-10-062-254-96

Query Match 88.9%; Score 24; DB 12; Length 176;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||
Db 117 DHALK 121

RESULT 23
US-10-062-254-88
Sequence 88, Application US/10062254
Patent No. US20020138882A1
GENERAL INFORMATION:
APPLICANT: Cahoon, Edgar B
APPLICANT: Cahoon, Rebecca E
APPLICANT: Falco, Saverio Carl
APPLICANT: Fang, Yiwen
APPLICANT: Hantke, Sabine S.
APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism

FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/062,254
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 09/630,346
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/146511
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/156006
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/156899
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 60/157287
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/169767
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171054
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/172958
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/171515
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 60/173535
NUMBER OF SEQ ID NOS: 375
SOFTWARE: Microsoft Office 97
SEQ ID NO 88
LENGTH: 222
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: UNSURE
LOCATION: (4)
NAME/KEY: UNSURE
LOCATION: (9)
NAME/KEY: UNSURE
LOCATION: (18)
NAME/KEY: UNSURE
LOCATION: (36)
NAME/KEY: UNSURE
LOCATION: (29)
NAME/KEY: UNSURE
LOCATION: (52)
US-10-062-254-88

Query Match 88.9%; Score 24; DB 12; Length 222;
Best Local Similarity 80.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
|||
Db 109 DHALK 113

RESULT 24
US-10-062-254-90
Sequence 90, Application US/10062254
Patent No. US20020138882A1
GENERAL INFORMATION:
APPLICANT: Cahoon, Edgar B
APPLICANT: Cahoon, Rebecca E
APPLICANT: Falco, Saverio Carl
APPLICANT: Fang, Yiwen
APPLICANT: Hantke, Sabine S.
APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun

;; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
;; FILE REFERENCE:
;; CURRENT FILING DATE: 2002-02-01
;; CURRENT APPLICATION NUMBER: US/10/062,254
;; PRIOR FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: 09/630,346
;; PRIOR FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: 60/146511
;; PRIOR FILING DATE: 1999-07-30
;; PRIOR APPLICATION NUMBER: 60/156006
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/156899
;; PRIOR FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: 60/157287
;; PRIOR FILING DATE: 1999-10-01
;; PRIOR APPLICATION NUMBER: 60/169767
;; PRIOR FILING DATE: 1999-12-09
;; PRIOR APPLICATION NUMBER: 60/171054
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: 60/172958
;; PRIOR FILING DATE: 1999-12-21
;; PRIOR APPLICATION NUMBER: 60/171515
;; PRIOR FILING DATE: 1999-12-22
;; PRIOR APPLICATION NUMBER: 60/173535
;; PRIOR FILING DATE: 1999-12-29
;; NUMBER OF SEQ ID NOS: 375
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 90
;; LENGTH: 250
;; TYPE: PRT
;; ORGANISM: Zea mays
US-10-062-254-90

Query Match 88.9%; Score 24; DB 12; Length 250;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 240 DHALK 244

RESULT 25
US-10-062-254-100
; Sequence 100, Application US/10062254
; Patent No. US2002013882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30

;; PRIOR APPLICATION NUMBER: 60/157287
;; PRIOR FILING DATE: 1999-10-01
;; PRIOR APPLICATION NUMBER: 60/169767
;; PRIOR FILING DATE: 1999-12-09
;; PRIOR APPLICATION NUMBER: 60/171054
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: 60/172958
;; PRIOR FILING DATE: 1999-12-21
;; PRIOR APPLICATION NUMBER: 60/171515
;; PRIOR FILING DATE: 1999-12-22
;; PRIOR APPLICATION NUMBER: 60/173535
;; PRIOR FILING DATE: 1999-12-29
;; NUMBER OF SEQ ID NOS: 375
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 100
;; LENGTH: 295
;; TYPE: PRT
;; ORGANISM: Glycine max
;; FEATURE:
;; NAME/KEY: UNSURE
;; LOCATION: (256)
US-10-062-254-100

Query Match 88.9%; Score 24; DB 12; Length 295;
Best Local Similarity 80.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 246 DHALK 250

Search completed: June 20, 2003, 21:02:33
Job time : 13.8409 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 12.7273 Seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-14
Perfect score: 27
Sequence: 1 DRAVX 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR 73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27	100.0	199	2 S65071	cystatin - field m
2	27	100.0	205	2 T14386	cysteine proteinas
3	27	100.0	206	2 T29696	hypothetical prote
4	27	100.0	250	2 AD0948	conserved hypothet
5	27	100.0	351	2 H90178	phosphate regulato
6	27	100.0	407	2 S51448	hypothetical prote
7	27	100.0	495	2 C89778	hypothetical prote
8	27	100.0	508	2 G90830	hypothetical prote
9	27	100.0	508	2 D85688	unknown protein en
10	27	100.0	508	2 F64786	ybck protein - Esc
11	27	100.0	550	2 T10393	chitinase (EC 3.2.
12	27	100.0	581	2 E69322	dolichol-P-glucose
13	27	100.0	632	2 T48940	hypothetical prote
14	27	100.0	632	2 D98191	trkA-like protein
15	27	100.0	632	2 AF3095	nitric oxide reduc
16	27	100.0	723	2 E82353	DNA helicase II VC
17	27	100.0	818	2 T40491	hypothetical prote
18	27	100.0	1224	2 S70100	MSN5 protein - yea
19	27	100.0	1248	1 D42088	adenyate cyclase
20	26	96.3	146	2 F64650	conserved hypothet
21	26	96.3	167	2 A49694	nonheme ferritin -
22	26	96.3	167	2 D71914	nonheme iron-conta
23	26	96.3	202	2 G86302	Flr16.6 protein -
24	26	96.3	243	2 T09769	uridyate kinase (
25	26	96.3	270	2 H69074	formylmethanofuran
26	26	96.3	270	2 S57457	mannose-6-phosphat
27	26	96.3	465	2 G75104	H+-transporting tw
28	26	96.3	467	2 D84938	glu-tRNA amidotran
29	26	96.3	485	2 D70142	

30	26	96.3	525	2 D90048	hypothetical prote
31	26	96.3	712	2 S50969	probable membrane
32	26	96.3	1244	2 T13068	hypothetical prote
33	26	96.3	2492	1 MNWVD	nonstructural poly
34	26	96.3	2492	1 A44213	nonstructural poly
35	26	96.3	2492	1 C44213	nonstructural poly
36	24	88.9	88	2 F64631	conserved hypothet
37	24	88.9	102	2 A69947	PTS system, cellob
38	24	88.9	110	2 D84244	inosine-5', monopho
39	24	88.9	136	2 AF1748	hypothetical prote
40	24	88.9	148	2 T05879	hypothetical prote
41	24	88.9	151	2 F81445	bacterioferritin c
42	24	88.9	153	2 A33090	conglutin delta pr
43	24	88.9	158	2 G69753	two-component sens
44	24	88.9	193	2 A49704	transcription init
45	24	88.9	206	2 F84087	hypothetical prote
46	24	88.9	215	2 AE2892	hypothetical prote
47	24	88.9	216	2 F84292	hypothetical prote
48	24	88.9	218	2 E97497	22K outer membrane
49	24	88.9	218	2 AB2716	outer membrane pro
50	24	88.9	229	1 B43685	nonstructural prot
51	24	88.9	236	2 D86803	prophage p13 prote
52	24	88.9	238	2 T12086	hypothetical prote
53	24	88.9	243	2 H82253	ribonuclease T VCI
54	24	88.9	271	2 D86405	unknown protein [i
55	24	88.9	273	2 F91083	hypothetical membr
56	24	88.9	278	2 G84023	phosphate ABC tran
57	24	88.9	281	2 T02369	hypothetical prote
58	24	88.9	281	2 E96679	hypothetical prote
59	24	88.9	284	2 B27407	tropomyosin alpha
60	24	88.9	286	2 AB4141	hypothetical prote
61	24	88.9	297	2 G87187	phosphoribosylamin
62	24	88.9	298	2 T12084	hypothetical prote
63	24	88.9	299	2 S72173	senescence marker
64	24	88.9	299	2 S34588	senescence marker
65	24	88.9	301	2 G85928	hypothetical prote
66	24	88.9	302	2 S75978	hypothetical prote
67	24	88.9	307	2 AG3414	dihydroadipicinat
68	24	88.9	316	2 T35006	beta-ketoacyl-acyl
69	24	88.9	327	2 H97667	hypothetical prote
70	24	88.9	333	2 G87154	DTDP-glucose 4,6-d
71	24	88.9	351	2 B69348	magnesium and coba
72	24	88.9	359	2 H64547	peptide methionine
73	24	88.9	360	2 H71828	probable rRNA (5-m
74	24	88.9	360	2 G46686	conserved hypothet
75	24	88.9	364	2 B36313	hypothetical 42K p

ALIGNMENTS

RESULT 1

S65071 cystatin - field mustard
N:Alternate names: cysteine proteinase inhibitor
C:Species: Brassica campestris (field mustard)
A:Variety: subsp. pekinensis
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Sep-1997
C:Accession: S65071
R:Lim, C.O.; Lee, S.I.; Chung, W.S.; Park, S.H.; Hwang, I.; Cho, M.J.
Plant Mol. Biol. 30, 373-379, 1996
A:Title: Characterization of a cDNA encoding cysteine proteinase inhibitor from Chi
A:Reference number: S65071; MUID:96178877; PMID:8616262
A:Accession: S65071
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-199 <LIM>
A:Cross-references: EMBL:L41355; NID:G762784; PID:G762785

Query Match 100.0%; Score 27; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
 |||||
 Db 126 DHAVK 130

RESULT 2

T14386

cysteine proteinase inhibitor BCPI-2 - turnip

C;Species: Brassica rapa (turnip)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T14386

R;Lim, C.O.; Beachy, R.N.; Cho, M.J.

submitted to the EMBL Data Library, March 1996

A;Description: Cysteine proteinase inhibitor cDNAs from Chinese cabbage.

A;Reference number: Z18021

A;Accession: T14386

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-205 <LM>

A;Cross-references: EMBL:U51119; NID:g1256423; PID:g1256424

A;Experimental source: subspecies pekinensis

C;Genetics:

A;Gene: BCPI-2

C;Keywords: cysteine proteinase inhibitor

Query Match 100.0%; Score 27; DB 2; Length 205;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
 |||||
 Db 132 DHAVK 136

RESULT 3

T29696

hypothetical protein F31A3.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Accession: T29696

R;Murray, J.; Le, T.T.

submitted to the EMBL Data Library, May 1996

A;Description: The sequence of C. elegans cosmid F31A3.

A;Reference number: Z20667

A;Accession: T29696

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-206 <MUT>

A;Cross-references: EMBL:U58742; PIDN:AAB36857.1; GSPDB:GN00028; CESP:F31A3.3

A;Experimental source: strain Bristol N2; clone F31A3

C;Genetics:

A;Gene: CESP:F31A3.3

A;Map position: X

A;Introns: 7/1; 67/3; 110/3; 138/3

Query Match 100.0%; Score 27; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
 |||||
 Db 179 DHAVK 183

RESULT 4

AD0948

conserved hypothetical protein STY3858 [imported] - Salmonella enterica subsp. enterica

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: This species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001

C;Accession: AD0948

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Steven

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica

A;Reference number: AB0502; PMID:11677608

A;Accession: AD0948

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-250 <PAB>

A;Cross-references: GB:AL513382; PIDN:CAD09606.1; PID:g16504717; GSPDB:GN00176

C;Genetics:

A;Gene: STY3858

Query Match 100.0%; Score 27; DB 2; Length 250;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
 |||||
 Db 202 DHAVK 206

RESULT 5

H90178

phosphate regulatory protein, probable [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: H90178

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.;

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Red

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: H90178

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-351 <KUR>

A;Cross-references: GB:AE006641; NID:gl3813501; PIDN:AAK40687.1; GSPDB:GN00155

C;Genetics:

A;Gene: SSO0356

Query Match 100.0%; Score 27; DB 2; Length 351;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
 |||||
 Db 264 DHAVK 268

RESULT 6

S51448

hypothetical protein YLR225c - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein L8083.10

C;Species: Saccharomyces cerevisiae

C;Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 19-Apr-2002

C;Accession: S51448

R;Hallsworth, K.

submitted to the EMBL Data Library, December 1994

A;Description: The sequence of S. cerevisiae cosmid 8083.

A;Reference number: S51443

A;Accession: S51448

A;Molecule type: DNA

A;Residues: 1-407 <HAL>

A;Cross-references: EMBL:U19027; NID:g609363; PIDN:AB67413.1; PID:g609373; GSPDB:

C;Genetics:

A;Gene: MIPS:YLR225C

A;Cross-references: SGD:S0004215

A;Map position: 12R

Query Match 100.0%; Score 27; DB 2; Length 407;
 Best Local Similarity 100.0%; Pred. No. 99;

QY 1 DHAVK 5
|||||
Db 411 DHAVK 415

RESULT 12
E69322
dolichol-P-glucose synthetase homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: E69322
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250, MUID:98049343, PMID:9389475
A:Accession: E69322
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-581 <KLE>
A:Cross-references: GB:AE001064; GB:AE000782; NID:G2689387; PIDN:AAB90655.1; PID:G265003

Query Match 100.0%; Score 27; DB 2; Length 581;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||||
Db 552 DHAVK 556

RESULT 13
T48940
hypothetical protein F14L2.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T48940
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.;
submitted to the Protein Sequence Database, April 2000
A:Reference number: 225008
A:Accession: T48940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-629 <JOR>
A:Cross-references: EMBL:AL353818; GSPDB:GN00061; ATSP:F14L2.150
A:Experimental source: cultivar Columbia; BAC clone F14L2
C:Genetics:
A:Gene: ATSP:F14L2.150
A:Map position: 3
A:Introns: 62/2; 91/1; 120/3; 200/3; 241/2; 266/3; 302/3; 361/3; 434/2; 476/3; 551/2; 60

Query Match 100.0%; Score 27; DB 2; Length 629;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||||
Db 134 DHAVK 138

RESULT 14
D98191
trkA-like protein (AF083948) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98191
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacteri
A:Reference number: A97359; PMID:11743194
A:Accession: D98191
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-632 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK9054.1; PID:gl5158850; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_961
A:Map position: linear chromosome

Query Match 100.0%; Score 27; DB 2; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||||
Db 30 DHAVK 34

RESULT 15
AF3095
nitric oxide reductase Nord protein [imported] - Agrobacterium tumefaciens (strain
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AF3095
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; M
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF3095
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-632 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL45180.1; PID:gl7742857; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: nord
A:Map position: linear chromosome

Query Match 100.0%; Score 27; DB 2; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||||
Db 30 DHAVK 34

RESULT 16
E82353
DNA helicase II VC0190 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: E82353
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson,
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sell
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82353
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-723 <HBI>
A:Cross-references: GB:AE004109; GB:AE003852; NID:G9654590; PIDN:AAF93366.1; GSPDB:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0190

A;Map position: 1
C;Superfamily: helicase II

Query Match 100.0%; Score 27; DB 2; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||
Db 482 DHAVK 486

RESULT 17

T40491

hypothetical protein SPBC4C3.06 - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40491
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, February 1998
A;Reference number: 221910
A;Accession: T40491
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-818 <WOO>
A;Cross-references: EMBL:AL021730; PIDN:CAA16828.1; GSPDB:GN00067; SPDB:SPBC4C3.06
A;Experimental source: strain 972h-; cosmid C4C3
C;Genetics:
A;Gene: SPDB:SPBC4C3.06
A;Map position: 2
A;Introns: 14/3

Query Match 100.0%; Score 27; DB 2; Length 818;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||
Db 612 DHAVK 616

RESULT 18

S70100

MSN5 protein - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein YDR335w
C;Species: Saccharomyces cerevisiae
C;Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 24-Nov-1999
C;Accession: S70100
R;Du, Z.
submitted to the EMBL Data Library, March 1996
A;Description: The sequence of S. cerevisiae cosmid 9651.
A;Reference number: S70098
A;Accession: S70100
A;Molecule type: DNA
A;Residues: 1-1224 <DUZ>
A;Cross-references: EMBL:U51032; NID:g1230659; PID:g1230662; MIPS:YDR335w
C;Genetics:
A;Gene: SGD:MSNS
A;Cross-references: SGD:S0002743; MIPS:YDR335w
A;Map position: 4R
C;Superfamily: Saccharomyces cerevisiae MSN5 protein
C;Keywords: transmembrane protein
F;158-174/Domain: transmembrane #status predicted <TM1>
F;186-202/Domain: transmembrane #status predicted <TM2>
F;869-885/Domain: transmembrane #status predicted <TM3>

Query Match 100.0%; Score 27; DB 2; Length 1224;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||
Db 68 DHAVK 72

RESULT 19

D42088

adenylate cyclase (EC 4.6.1.1), Ca2+/calmodulin-responsive - fruit fly (Drosophila)
N;Alternate names: rutabaga protein
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: D42088
R;Levin, L.R.; Han, P.L.; Hwang, P.M.; Feinstein, P.G.; Davis, R.L.; Reed, R.R.
Cell 68, 479-489, 1992

A;Title: The Drosophila learning and memory gene rutabaga encodes a Ca2+/calmodulin
A;Reference number: A42088; MUID:92154664; PMID:1739965
A;Accession: D42088
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-2248 <LEV>
A;Cross-references: GB:M81887; NID:gl58194; PIDN:AAA28844.1; PID:gl58195
A;Note: sequence extracted from NCBI backbone (NCBIP:82229)
C;Genetics:
A;Gene: FlyBase:rut

A;Cross-references: FlyBase:FBgn0003301
C;Superfamily: Drosophila adenylate cyclase; guanylate cyclase catalytic domain hom
C;Keywords: phosphorus-oxygen lyase; transmembrane protein
F;218-452/Domain: guanylate cyclase catalytic domain homology <GCC1>
F;910-1155/Domain: guanylate cyclase catalytic domain homology <GCC2>

Query Match 100.0%; Score 27; DB 1; Length 2248;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||
Db 2 DHAVK 6

RESULT 20

F64650

conserved hypothetical protein HPI046 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: F64650

R;Tomb, J.F.; White, O.; Kervage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; Mc
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watche
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frase
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: F64650
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-146 <TOM>
A;Cross-references: GB:AE000612; GB:AE000511; NID:g2314188; PIDN:AAD08091.1; PID:g2
C;Superfamily: nus operon 15K protein

Query Match 96.3%; Score 26; DB 2; Length 146;
Best Local Similarity 80.0%; Pred. No. 58;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||
Db 126 DHAIK 130

RESULT 21

A49694

nonheme ferritin - Helicobacter pylori (strains 26695, NCTC 11638, and others)
N;Alternate names: 19K adhesin; nonheme iron-containing ferritin
C;Species: Helicobacter pylori
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 20-Aug-1999
C;Accession: A49694; E64601; A41837
R;Frazier, B.A.; Pfeifer, J.D.; Russell, D.G.; Falk, P.; Olsen, A.N.; Hammar, M.; W

J. Bacteriol. 175, 966-972, 1993
A>Title: Paracrystalline inclusions of a novel ferritin containing nonheme iron, produce
A:Reference number: A49694; MUID:93163061; PMID:8432720
A:Accession: A49694
A:Molecule type: DNA; protein
A:Residues: 1-167 <FRA>
A:Cross-references: GB:S54729; NID:9265360; PIDN:AAB25329.1; PID:9265361
A:Experimental source: NCTC 11638
A:Notes: sequence extracted from NCBI backbone (NCBIN:124685, NCBIP:124686)
R:Tom, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frazer, C.
A>Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A:Reference number: A4520; MUID:97394467; PMID:9252185
A:Accession: E64601
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-38, 'A', 40-167 <TOM>
A:Cross-references: GB:AE000578; GB:AE000511; NID:92313759; PIDN:AAD07f12.1; PID:9231377
A:Experimental source: strain 28695
R:Doig, P.; Austin, J.W.; Kostzynska, M.; Trust, T.J.
J. Bacteriol. 174, 2539-2547, 1992
A>Title: Production of a conserved adhesin by the human gastroduodenal pathogen *Helicoba*
A:Reference number: A41837; MUID:92210497; PMID:1556073
A:Accession: A41837
A:Molecule type: protein
A:Residues: 1-28 <DOI>
A:Comment: This protein is characterized in reference A49694 as binding nonheme iron and
being loosely associated with and easily removed from the cell surface and as acting as a
C:Genetics:
A:Gene: HP0653
C:Function:
A:Description: involved in iron storage
A:Superfamily: ferritin
C:Keywords: homomultimer; iron storage

Query Match 96.3%; Score 26; DB 2; Length 167;
Best Local Similarity 80.0%; Pred. No. 67;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DHAVK 5
|||:|
Db 106 DHAIK 110

RESULT 22
D71914
nonheme iron-containing ferritin - *Helicobacter pylori* (strain J99)
C:Species: *Helicobacter pylori*
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Aug-1999
C:Accession: D71914
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: D71914
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-167 <ARN>
A:Cross-references: GB:AS001491; GB:AE001439; NID:94155127; PIDN:AAD06160.1; PID:9415513
A:Experimental source: strain J99
C:Genetics:
A:Gene: pfr
C:Superfamily: ferritin

Query Match 96.3%; Score 26; DB 2; Length 167;
Best Local Similarity 80.0%; Pred. No. 67;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DHAVK 5
|||:|
Db 106 DHAIK 110

RESULT 23

G86302
F17F16.6 protein - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001
C:Accession: G86302
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kzi
C.A.; Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal
ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86302
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-202 <STO>
A:Cross-references: GB:AE005172; NID:9954732; PIDN:AAG09085.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 96.3%; Score 26; DB 2; Length 202;
Best Local Similarity 80.0%; Pred. No. 82;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DHAVK 5
|||:|
Db 52 DHAIK 56

RESULT 24

T09769
uridylylate kinase (EC 2.7.4.-) - *Mycoplasma genitalium*
N:Alternate names: mukB suppressor protein emba homolog
C:Species: *Mycoplasma genitalium*
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Jun-2000
C:Accession: T09769; I64247
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann,
M.; Fuhrmann, J.L.; Nguyen, D.T.; Utterback, T.; Saudek, D.M.; Phillips, C.A.; Me
submitted to the EMBL Data Library, October 1998
A:Reference number: Z16818
A:Accession: T09769
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-243 <FRA>
A:Cross-references: EMBL:U39725; NID:93845017; PID:93845027
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann,
M.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merri
C.A.; Venter, J.C.
Science 270, 397-403, 1995
A>Title: The minimal gene complement of *Mycoplasma genitalium*.
A:Reference number: A64200; MUID:96026346; PMID:7569993
A:Accession: I64247
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 'M', 115-243 <TIGR>
A:Cross-references: GB:U39730; GB:LA3967; NID:91046149; PIDN:AAB01624.1; PID:910461
A:Experimental source: strain G-37
C:Genetics:
A:Gene: MG434
A:Genetic code: SGC3
C:Superfamily: uridine 5'-monophosphate kinase
C:Keywords: phosphotransferase

Query Match 96.3%; Score 26; DB 2; Length 243;
Best Local Similarity 80.0%; Pred. No. 1.e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:
DB 191 DRAIK 195

RESULT 25

H69074
formylmethanofuran dehydrogenase (EC 1.2.99.5) (tungsten) chain C - Methanobacterium the
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
C;Accession: H69074
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: H69074
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-270 <MTH>
A;Cross-references: GB:AE000916; GB:AE000666; NID:G2622674; PIDN:AAB86032.1; PID:G262268
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1558
C;Superfamily: formylmethanofuran dehydrogenase chain C
C;Keywords: iron-sulfur protein; metalloprotein; oxidoreductase; tungsten

Query Match 96.3%; Score 26; DB 2; Length 270;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:
DB 248 DRAIK 252

Search completed: June 20, 2003, 20:59:01
Job time : 15.7273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 ; Search time 6.02273 Seconds
(without alignments)
34.433 Million cell updates/sec

Title: US-10-105-008-14

Perfect score: 27

Sequence: 1 DRAVK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	508	YBCK_ECOLI	P77698 escherichia
2	27	100.0	550	CHIT_NFVOP	O10363 orgyia pseu
3	27	100.0	1113	CARB_CORGL	P58939 corynebacte
4	27	100.0	1224	MSN5_YEAST	P52918 saccharomyc
5	27	100.0	1227	LAP4_HUMAN	P51826 homo sapien
6	27	100.0	2248	CYAI_DROME	P32870 drosophila
7	26	96.3	146	YA46_HELPY	O25687 helicobacte
8	26	96.3	167	FTN_HELPY	Q92111 helicobacte
9	26	96.3	167	FTN_HELPY	P52093 helicobacte
10	26	96.3	243	PYB8_MYCGE	P47672 mycoplasma
11	26	96.3	270	FWDC_METTH	O27600 methanobact
12	26	96.3	270	FWDC_METTH	Q59579 methanobact
13	26	96.3	270	FWDC_METTH	O74031 methanobact
14	26	96.3	350	WNT2_BRARE	Q92048 brachydanio
15	26	96.3	467	FLI1_BUCAL	P57178 buchnera ap
16	26	96.3	485	GATE_BORBU	O51316 borrelia bu
17	26	96.3	712	FRB6_YEAST	Q12473 saccharomyc
18	26	96.3	1072	CARB_THETIN	Q81801 thermomancr
19	26	96.3	1244	YP83_CABEL	Q09441 caenorhabdi
20	26	96.3	2485	POLN_ERVW3	P36327 venezuelan
21	26	96.3	2492	POLN_ERVVP	P36328 venezuelan
22	26	96.3	2492	POLN_ERVVT	P27282 venezuelan
23	24	88.9	112	GLNB_RHET	O54053 rhizobium e
24	24	88.9	158	YC8M_BACSU	P42245 bacillus su
25	24	88.9	193	YP8H_PSEAE	Q06198 pseudomonas
26	24	88.9	224	RNT_VIBCH	Q9kt97 vibrio chol
27	24	88.9	229	RAD8_THEAC	Q9bjd3 thermoplas
28	24	88.9	229	VU_P14HA	P21739 human parai
29	24	88.9	284	TPM1_RAT	P04692 rattus norv
30	24	88.9	297	PUR7_MYCLE	O08361 mycobacteri
31	24	88.9	299	SM30_MOUSE	Q64374 mus musculu
32	24	88.9	299	SM30_MOUSE	Q03336 rattus norv
33	24	88.9	347	RNHL_DROME	Q9vpp5 drosophila

34	24	88.9	359	1	MSAB_HELPY	O25011 h peptide m
35	24	88.9	360	1	TRMU_HELPY	Q9ajq0 helicobacte
36	24	88.9	360	1	TRMU_HELPY	O25893 helicobacte
37	24	88.9	382	1	ADH2_ZYMO	P06758 zymomonas m
38	24	88.9	399	1	R3PP_P14HA	P22044 human parai
39	24	88.9	404	1	ISCS_VIBCH	Q9kt92 vibrio chol
40	24	88.9	412	1	DAD3_RHILO	Q981X2 rhizobium l
41	24	88.9	422	1	ADHF_SCHPO	Q09669 schizosacch
42	24	88.9	455	1	HYIN_PSESS	P06618 pseudomonas
43	24	88.9	464	1	RMV7_ECOLI	P37741 escherichia
44	24	88.9	499	1	KPY1_TRYBB	P30615 trypanosoma
45	24	88.9	499	1	KPY2_TRYBB	P30616 trypanosoma
46	24	88.9	508	1	SYM_RICPR	Q98cp0 rickettsia
47	24	88.9	601	1	SG2_RANRI	P30945 rana ridibu
48	24	88.9	623	1	HCVE_EURCA	P02242 eurytelma c
49	24	88.9	654	1	CNIC_MOUSE	P37741 escherichia
50	24	88.9	662	1	TLPA_BACSU	Q64338 mus musculu
51	24	88.9	666	1	FLID_VIBCH	P39216 bacillus au
52	24	88.9	681	1	MF10_HUMAN	O00566 homo sapien
53	24	88.9	709	1	CNIC_HUMAN	Q14123 homo sapien
54	24	88.9	720	1	PAL1_SOLTU	P31425 solanum tub
55	24	88.9	721	1	PAL5_LYCER	P26600 lycopersico
56	24	88.9	740	1	POLB_MAIZE	P15718 zea mays (m
57	24	88.9	768	1	CNIC_RAT	Q63421 rattus norv
58	24	88.9	911	1	CIAP_BACTU	P96315 bacillus th
59	24	88.9	968	1	ST10_HUMAN	O94804 homo sapien
60	24	88.9	1009	1	DPOL_HSVSA	P24907 herpesvirus
61	24	88.9	1052	1	BULB_MOUSE	Q92180 mus musculu
62	24	88.9	1077	1	UPPS_HUMAN	Q96ru2 homo sapien
63	24	88.9	1155	1	CIAB_BACTK	P08578 bacillus th
64	24	88.9	1272	1	FMK2_MOUSE	O51112 mus musculu
65	24	88.9	1311	1	FMK2_MOUSE	P51816 homo sapien
66	24	88.9	2014	1	YU07_YEAST	P39526 saccharomyc
67	24	88.9	2431	1	POLN_SFV	P08411 senliki for
68	24	88.9	2514	1	POLN_ONNVG	P13886 o'nyong-nyo
69	23	85.2	87	1	ACBP_FRIAG	O22643 fritillaria
70	23	85.2	126	1	DHA5_BOVIN	P52476 bos taurus
71	23	85.2	144	1	R12C_ARATH	Q96kz3 arabidopsis
72	23	85.2	150	1	RL13_CHLMU	Q9pkz3 chlamydia m
73	23	85.2	150	1	RL13_CHLYR	O84127 chlamydia t
74	23	85.2	151	1	DUT_YERPE	Q88jp5 yersinia pe
75	23	85.2	197	1	GRPE_LACSK	O87776 lactobacill

ALIGNMENTS

RESULT 1
YBCK_ECOLI
ID YBCK_ECOLI STANDARD; PRT; 508 AA.
AC P77698;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybck.
GN YBCK OR B0544
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,

RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
 RA Nanath A., Oefner P., Roberts D., Schramm S., Davis R.W.,
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000160; AAC73645.1; -;
 DR EMBL; U82598; BAB40741.1; -;
 DR EcoGene; EGI3627; ybck.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 508 AA; 57526 MW; F85C07DBP45105CD CRC64;

Query Match 100.0%; Score 27; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DHAVK 5
 Db 219 DHAVK 223

RESULT 2
 CHIT NPVOV STANDARD; PRT; 550 AA.
 AC O1033;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Probable endochitinase precursor (EC 3.2.1.14).
 OS Orygia pseudotugata multicapsid polyhedrosis virus (OpMPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=164623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97271300; PubMed=9126251;
 RX Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
 RA Rohrmann G.F.;
 RT "The sequence of the Orygia pseudotugata multinucleocapsid nuclear
 RT polyhedrosis virus genome.";
 RL Virology 229:381-399 (1997).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
 CC acetyl-D-glucosamine polymers of chitin.
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U75930; AAC59123.1; -;
 DR HSP; P07254; 1CTN.
 DR InterPro; IPR001579; Chitinase_18/2.
 DR InterPro; IPR000886; ER target.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR InterPro; IPR000601; PKD domain.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR ProDom; PD000471; Glyco_hydro_18; 1.
 DR SMART; SM00089; PKD; 1.
 DR PROSITE; PS00014; ER TARGET; 1.
 DR PROSITE; PS01095; CHITINASE 18; 1.
 KW Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein;

KW Endoplasmic reticulum.
 FT SIGNAL 1 16
 FT CHAIN 17 550
 FT ACT SITE 304 304
 FT CARBOHYD 146 146
 FT CARBOHYD 172 172
 FT CARBOHYD 344 344
 FT SITE 547 550
 SQ SEQUENCE 550 AA; 60733 MW; 77947F5CF0E07BD CRC64;

Query Match 100.0%; Score 27; DB 1; Length 550;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DHAVK 5
 Db 411 DHAVK 415

RESULT 3
 CARB CORGL STANDARD; PRT; 1113 AA.
 AC P58939;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
 DE phosphate synthetase ammonia chain).
 GN CARB OR CGL1609.
 OS Corynebacterium glutamicum (Brevibacterium flavyum).
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
 CC phosphate + L-glutamate + carbamoyl phosphate.
 CC -1- COFACTOR: Binds three manganese ions (By similarity).
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
 CC promotes the hydrolysis of glutamine to ammonia, which is used by
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP005279; BAB99002.1; -;
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 KW ATP-binding; Manganese.
 FT DOMAIN 1 407
 FT DOMAIN 408 565
 FT DOMAIN 566 967
 FT DOMAIN 968 1113
 FT REPEAT 1 565
 FT REPEAT 566 1113
 FT NP_BIND 158 215
 FT NP_BIND 308 358
 FT METAL 290 290
 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 OLIGOMERIZATION DOMAIN.
 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 ALLOSTERIC DOMAIN.
 ATP (POTENTIAL).
 ATP (POTENTIAL).
 MANGANESE 1 (BY SIMILARITY).

```
FT METAL 304 304 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 306 306 MANGANESE 2 (BY SIMILARITY).
FT METAL 845 845 MANGANESE 3 (BY SIMILARITY).
FT METAL 857 857 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1113 AA; 120037 MW; 85D2DA1A4DF66CA CRC64;

Query Match 100.0%; Score 27; DB 1; Length 1113;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
Db 1108 DHAVK 1112

RESULT 4
MSNS_YEAST STANDARD; PRT; 1224 AA.
AC P52918;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MSN5 protein.
DE MSN5 OR YDR335W OR D9651.5.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Alepuz P.M., Eatruch F.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifkin L., Riles L., Taich A., Trevisan E., Vignati D.,
RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; X93302; CAA63705.1; -
DR EMBL; U51032; AAB64771.1; -
DR SGD; S0002743; MSN5 D -> G (IN REF. 1).
FT CONFLICT 886 886 S -> R (IN REF. 1).
FT CONFLICT 1011 1011
SQ SEQUENCE 1224 AA; 142117 MW; C47A6767CBA4486 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 1224;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
Db 68 DHAVK 72

RESULT 5
LAF4_HUMAN STANDARD; PRT; 1227 AA.
ID P51826;
AC
```

```
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE LAF-4 protein (Lymphoid nuclear protein related to AP4).
GN LAF4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96141096; PubMed=8555498;
RA Ma C., Staudt L.M.;
RT "LAF-4 encodes a lymphoid nuclear protein with transactivation
RT potential that is homologous to AP-4, the gene fused to MLL in
RT t(4;11) leukemias."
RL Blood 87:734-745(1996).
CC -!- FUNCTION: PUTATIVE TRANSCRIPTION ACTIVATOR THAT MAY FUNCTION IN
CC LYMPOID DEVELOPMENT AND ONCOGENESIS. BINDS, IN VITRO, TO
CC DOUBLE-STRANDED DNA.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN LYMPHOID TISSUES,
CC HIGHEST LEVELS BEING FOUND IN THE THYMUS.
CC -!- SIMILARITY: BELONGS TO THE AP4 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; U34360; AAA98763.1; -
DR TRANSFAC; T04866; -
DR Genew; HGNC:6473; LAF4.
DR MIM; 601464; -
KW Nuclear protein; Transcription regulation; Activator; DNA-binding.
FT DOMAIN 413 419 POLY-SER.
FT DOMAIN 422 432 POLY-SER.
FT DOMAIN 440 445 POLY-SER.
FT DOMAIN 670 679 POLY-SER.
SQ SEQUENCE 1227 AA; 133734 MW; 634B896FD7E9BBE7 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 1227;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
Db 1082 DHAVK 1086

RESULT 6
CYAL DROME STANDARD; PRT; 2248 AA.
AC P32870;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ca2+/calmodulin-responsive adenylate cyclase (EC 4.6.1.1) (ATP
DE pyrophosphate-lyase) (Rutabaga protein).
GN RUT.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE=92154664; PubMed=1739965;
RA Levin L.R., Han P.-L., Hwang P.M., Feinstein P.G., Davis R.L.,
```

RA Read R.R.;
RT Ca2+/Calmodulin-responsive adenylyl cyclase.";
RL Cell 68:479-489(1992).
CC -1- FUNCTION: THIS A MEMBRANE-BOUND, CALMODULIN-SENSITIVE ADENYLYL
CYCLASE. INACTIVATION OF THIS CYCLASE LEADS TO A LEARNING AND
MEMORY DEFECT.
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY CA(2+)/CALMODULIN AND G PROTEIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: MUSHROOM BODIES OF THE FLY BRAIN.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL; M81887; AAA28844.1; -;
CC PIR; D42088; D42088.
CC HSSP; P19754; 1AWK.
CC FlyBase; FBgn0003301; rut.
CC InterPro; IPR001054; G_cyclase.
CC Pfam; PF00211; guanylate_cyc; 2.
CC SMART; SM00044; CYC; 2.
CC PROSITE; PS00452; GUANYLATE_CYCLASES_1; 2.
CC PROSITE; PS00452; GUANYLATE_CYCLASES_2; 2.
CC Lyase; CAMP synthesis; Transmembrane; Glycoprotein; Repeat.
KW DOMAIN 1 41
FT TRANSMEM 42 60
FT TRANSMEM 65 84
FT TRANSMEM 101 115
FT TRANSMEM 122 142
FT TRANSMEM 152 174
FT TRANSMEM 186 206
FT DOMAIN 207 705
FT TRANSMEM 706 726
FT TRANSMEM 730 750
FT TRANSMEM 770 791
FT DOMAIN 792 813
FT TRANSMEM 814 834
FT TRANSMEM 842 867
FT TRANSMEM 868 888
FT DOMAIN 889 2248
FT TRANSMEM 318 334
FT DOMAIN 1013 1029
FT DOMAIN 515 530
FT DOMAIN 569 602
FT DOMAIN 1278 1297
FT DOMAIN 1767 1810
FT DOMAIN 2025 2040
FT DOMAIN 2200 2241
FT CARBOHYD 800 800
FT CARBOHYD 807 807
FT MUTAGEN 1026 1026
SQ SEQUENCE 2248 AA; 248899 MW; E459C718BE018668 CRC64;
Query Match 100.0%; Score 27; DB 1; Length 2248;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 2 DHAVK 6

RESULT 7
YA46_HELPY

YA46_HELPY STANDARD; PRT; 146 AA.
Q25687;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein HP1046.
GN HP1046.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97334467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Cocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0090 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL; AE000612; AAD08091.1; -;
CC TIGR; HP1046;
CC InterPro; IPR003728; DUF150.
CC Pfam; PF02576; DUF150; 1.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 146 AA; 16543 MW; 9DD5A08C7B759BD4 CRC64;
SQ
Query Match 96.3%; Score 26; DB 1; Length 146;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
DB 126 DHAVK 130

RESULT 8
FTN_HELPY
ID FTN_HELPY STANDARD; PRT; 167 AA.
AC Q25687;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonheme iron-containing ferritin.
GN PPR OR JHP0598.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJongh B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,

RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RL gastric pathogen *Helicobacter pylori*."
RL Nature 397:176-180(1999).
CC -|- FUNCTION: IRON-STORAGE PROTEIN (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: BELONGS TO THE FERRITIN FAMILY. PROKARYOTIC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001491; AAD06160.1; -;
DR HSSP; P23887; LEUM.
DR InterPro; IPR001519; Ferritin.
DR Pfam; PF00210; ferritin; 1.
DR TIGR; HP0653; -;
DR TIGR; HP0653; -;
DR InterPro; IPR001519; Ferritin.
DR Pfam; PF00210; ferritin; 1.
DR Iron storage; Complete proteome.
FT METAL 17 17 IRON (BY SIMILARITY).
SQ SEQUENCE 167 AA; 19314 MW; D1887P3P2CAD9CFC CRC64;

Query Match 96.3%; Score 26; DB 1; Length 167;
Best Local Similarity 80.0%; Pred. No. 21;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
DB 106 DHAIK 110

RESULT 9
FTN_HELPY
ID FTN_HELPY STANDARD; PRT; 167 AA.
AC P52093;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonheme iron-containing ferritin.
DE PFR OR HP0653.
GN *Helicobacter pylori* (Campylobacter pylori).
OS *Helicobacter pylori* (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
OC *Helicobacter*.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=CCUG 17874 / NCTC 11638;
RX MEDLINE=93163061; PubMed=8432720;
RA Frazer B.A., Pfeiffer J.D., Russell D.G., Falk P., Olesen A.N.,
RA Hammar M., Westblom T.U., Normark S.J.;
RT "Paracrystalline inclusions of a novel ferritin containing nonheme
RT iron, produced by the human gastric pathogen *Helicobacter pylori*:
RT evidence for a third class of ferritins."
RL J. Bacteriol. 175:966-972(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
RT *pylori*."
RL Nature 388:539-547(1997).
CC -|- FUNCTION: IRON-STORAGE PROTEIN.

CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO THE FERRITIN FAMILY. PROKARYOTIC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S54729; AAB25329.1; -;
DR EMBL; AE000578; AAD07712.1; -;
DR HSSP; P23887; LEUM.
DR TIGR; HP0653; -;
DR TIGR; HP0653; -;
DR InterPro; IPR001519; Ferritin.
DR Pfam; PF00210; ferritin; 1.
DR Iron storage; Complete proteome.
FT METAL 17 17 IRON (BY SIMILARITY).
FT CONFLICT 39 39 A -> S (IN REF. 1).
SQ SEQUENCE 167 AA; 19286 MW; 2829A4EC2AE616D4 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 167;
Best Local Similarity 80.0%; Pred. No. 21;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
DB 106 DHAIK 110

RESULT 10
PYRH_MYCGE
ID PYRH_MYCGE STANDARD; PRT; 243 AA.
AC P47672;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative uridylyate kinase (EC 2.7.4.-) (UK) (Uridine monophosphate
DE kinase) (UMP kinase).
GN PYRH OR SMBA OR MG434.
OS *Mycoplasma genitalium*.
OC Bacteria; Firmicutes; Mollicutes; *Mycoplasma* sp.;
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kervage A.R., Sutton G., Welley J.M.,
RA Fritchman J.L., Weidman J.P., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of *Mycoplasma genitalium*."
RL Science 270:397-403(1995).
CC -|- FUNCTION: URIDINE MONOPHOSPHATE KINASE (BY SIMILARITY).
CC -|- CATALYTIC ACTIVITY: ATP + UMP = ADP + UDP.
CC -|- SIMILARITY: TO OTHER UMP KINASE; SOME, TO ASPARTOKINASES AND TO
CC GLUTAMATE KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39725; AAC72455.1; -;
DR TIGR; MG434; -;
DR InterPro; IPR001048; Aa kinase.
DR Pfam; PF00696; aakinese; 1.

KW Transferase; Kinase; Pyrimidine biosynthesis; Complete proteome.
SQ SEQUENCE 243 AA; 26734 MW; 675AC6240FA7B65 CRC64;
Query Match 96.3%; Score 26; DB 1; Length 243;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
Db 191 DHAIK 195
RESULT 11
FWDC_METTH STANDARD; PRT; 270 AA.
AC O27600;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tungsten-containing formylmethanofuran dehydrogenase II subunit C
DE (EC 1.2.99.5).
GN FWDC OR MTH1558.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang R., Keagle P., Lum W., Pochier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- FUNCTION: CATALYZES THE REVERSIBLE OXIDATION OF CO2 AND
CC METHANOFURAN (MFR) TO N-FORMYLMETHANOFURAN (CHO-MFR). CAN ONLY
CC OXIDISE FORMYLMETHANOFURAN. THIS ENZYME IS OXYGEN-LABILE.
CC -1- CATALYTIC ACTIVITY: Formylmethanofuran + H(2)O + acceptor = CO(2)
CC + methanofuran + reduced acceptor.
CC -1- COFACTOR: TUNGSTEN.
CC -1- ENZYME REGULATION: NOT INACTIVATED BY CYANIDE.
CC -1- PATHWAY: FIRST STEP IN METHANOGENESIS.
CC -1- SUBUNIT: THIS ENZYME IS COMPOSED OF SEVEN SUBUNITS FWDA (65 kDa),
CC FWDB (53 kDa), FWDC (31 kDa), FWDD (15 kDa), FWDE, FWDF, AND FWDG.
CC -1- INDUCTION: BY GROWTH ON TUNGSTEN OR MOLYBDENUM UNDER ANAEROBIC
CC CONDITIONS.
CC -1- SIMILARITY: BELONGS TO THE FWDC/FMDC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AE000916; AAB86032.1; -
CC InterPro; IPR002489; DUF14.
CC Pfam; PF01493; DUF14; 1.
CC OXidoreductase; Tungsten; Methanogenesis; Repeat; Complete proteome.
CC DOMAIN 80 , 213 X-[IL]-X-[IV]-X-G.
CC REPEAT 80 92 1.
CC REPEAT 99 111 2.
CC REPEAT 118 130 3.
CC REPEAT 144 156 4.
CC REPEAT 163 175 5.
FT REPEAT 80 92 1.
FT REPEAT 99 111 2.
FT REPEAT 118 130 3.
FT REPEAT 144 156 4.
FT REPEAT 163 175 5.
Query Match 96.3%; Score 26; DB 1; Length 270;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DHAVK 5
Db 248 DHAIK 252
RESULT 12
FWDC_METTH STANDARD; PRT; 270 AA.
AC Q59579; O08493;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tungsten-containing formylmethanofuran dehydrogenase II subunit C
DE (EC 1.2.99.5).
GN FWDC.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=79929;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96163477; PubMed=8575452;
RA Hochheiner A., Schmitz R.A., Thauer R.K., Hedderich R.,
RT "The tungsten formylmethanofuran dehydrogenase from Methanobacterium
RT thermoautotrophicum contains sequence motifs characteristic for
RL Eur. J. Biochem. 234:910-920(1995).
CC -1- FUNCTION: CATALYZES THE REVERSIBLE OXIDATION OF CO2 AND
CC METHANOFURAN (MFR) TO N-FORMYLMETHANOFURAN (CHO-MFR). CAN ONLY
CC OXIDISE FORMYLMETHANOFURAN. THIS ENZYME IS OXYGEN-LABILE.
CC -1- CATALYTIC ACTIVITY: Formylmethanofuran + H(2)O + acceptor = CO(2)
CC + methanofuran + reduced acceptor.
CC -1- COFACTOR: TUNGSTEN.
CC -1- PATHWAY: FIRST STEP IN METHANOGENESIS.
CC -1- SUBUNIT: THIS ENZYME IS COMPOSED OF SEVEN SUBUNITS FWDA (65 kDa),
CC FWDB (53 kDa), FWDC (31 kDa), FWDD (15 kDa), FWDE, FWDF, AND FWDG
CC (BY SIMILARITY).
CC -1- INDUCTION: BY GROWTH ON TUNGSTEN OR MOLYBDENUM UNDER ANAEROBIC
CC CONDITIONS.
CC -1- SIMILARITY: BELONGS TO THE FWDC/FMDC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; X87970; CAA61214.1; -
CC InterPro; IPR002489; DUF14.
CC Pfam; PF01493; DUF14; 1.
CC OXidoreductase; Tungsten; Methanogenesis; Repeat.
CC DOMAIN 80 213 7 X-13 AA REPEATS OF [GW]-X-X-M-X-X-G-
CC X-[IL]-X-[IV]-X-G.
CC REPEAT 80 92 1.
CC REPEAT 99 111 2.
CC REPEAT 118 130 3.
CC REPEAT 144 156 4.
CC REPEAT 163 175 5.
CC REPEAT 182 194 6.
CC REPEAT 201 213 7.
SQ SEQUENCE 270 AA; 28554 MW; E0A369D2ACFECA6F CRC64;
Query Match 96.3%; Score 26; DB 1; Length 270;

Best Local Similarity 80.0%; Pred. No. 35;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:|
Db 248 DHAIK 252

RESULT 13
FWDC_METHO STANDARD; PRT; 270 AA.
AC O74031;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tungsten-containing formylmethanofuran dehydrogenase II subunit C
DE (ZC1.2.99.5).
GN FWDC.
OS Methanobacterium wolfei.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=145261;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99035764; PubMed=9818358;
RX Hochheiner A., Hedderich R., Thauer R.K.;
RT "The formylmethanofuran dehydrogenase isoenzymes in Methanobacterium
RT wolfei and Methanobacterium thermoautotrophicum: induction of the
RT molybdenum isoenzyme by molybdate and constitutive synthesis of the
RT tungsten isoenzyme.";
RL Arch. Microbiol. 170:389-393(1998).
CC -1- FUNCTION: CATALYZES THE REVERSIBLE OXIDATION OF CO2 AND
CC METHANOFURAN (MF) TO N-FORMYLMETHANOFURAN (CHO-MF). THIS ENZYME
CC IS OXYGEN-LABILE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Formylmethanofuran + H(2)O + acceptor = CO(2)
CC + methanofuran + reduced acceptor.
CC -1- COFACTOR: TUNGSTEN.
CC -1- PATHWAY: FIRST STEP IN METHANOGENESIS.
CC -1- SUBUNIT: THIS ENZYME IS COMPOSED OF SIX SUBUNITS FWDA, FWDB, FWDC,
CC FWDD, FWDE, AND FWDF (BY SIMILARITY).
CC -1- INDUCTION: BY FWDC OR TUNGSTEN OR MOLYBDENUM UNDER ANAEROBIC
CC CONDITIONS.
CC -1- SIMILARITY: BELONGS TO THE FWDC/FWDC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ009688; CAA08784.1;
CC InterPro; IPR002489; DUF14.
CC Pfam; PF01493; DUF14; 1.
KW Oxidoreductase; Tungsten; Methanogenesis; Repeat.
FT DOMAIN 80 213
FT X-[IL]-X-[IV]-X-G.
FT REPEAT 80 92
FT REPEAT 99 111
FT REPEAT 118 130
FT REPEAT 144 156
FT REPEAT 163 175
FT REPEAT 182 194
FT REPEAT 201 213
SQ SEQUENCE 270 AA; 28660 MW; BB1061CBDF53061B CRC64;

Query Match 96.3%; Score 26; DB 1; Length 270;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:|

Db 248 DHAIK 252

RESULT 14
WNT2_BRARE STANDARD; PRT; 350 AA.
AC O92048;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Wnt-2 protein precursor.
GN WNT2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
RC TISSUE=Embryo;
RA Blader P., Straehle U., Ingham P.W.;
RT "Three Wnt genes expressed in a wide variety of tissues during
RT development of the zebrafish, Danio rerio: developmental and
RT evolutionary perspectives.";
RL Dev. Genes Evol. 206:3-13(1996).
CC -1- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN
CC TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A
CC SIGNALING MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE
CC REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL
CC DIAMETERS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Possibly secreted and associates with the
CC extracellular matrix.
CC -1- DEVELOPMENTAL STAGE: First detected in the developing viscera at
CC 30 h. Expression is gone by 5 days.
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51266; AAA96517.1;
CC ZFIN; ZDB-GENE-980526-416; wnt2.
CC InterPro; IPR000970; Wnt_gthfactor.
CC Pfam; PF00110; wnt; 1.
CC PRINTS; PR01349; WNTPROTEIN.
CC SMART; SM00097; WNT1; 1.
CC PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein; Signal.
FT SIGNAL 1 25
FT POTENTIAL.
FT CHAIN 26 350
FT WNT-2 PROTEIN.
FT CARBOHYD 132 132
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 293 293
FT N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 350 AA; 139789 MW; F6268DADE9361C6C CRC64;

Query Match 96.3%; Score 26; DB 1; Length 350;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:|
Db 160 DHAIK 164

RESULT 15
FLII_BUCAI STANDARD; PRT; 467 AA.
ID FLII_BUCAI
AC P57178;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Flagellum-specific ATP synthase (EC 3.6.3.14).
 GN FLII OR BU076.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tokyo 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- FUNCTION: PROBABLE CATALYTIC SUBUNIT OF A PROTEIN TRANSLOCASE FOR
 CC FLAGELLUM-SPECIFIC EXPORT, OR A PROTON TRANSLOCASE INVOLVED IN
 CC LOCAL CIRCUITS AT THE FLAGELLUM. MAY BE INVOLVED IN A SPECIALIZED
 CC PROTEIN EXPORT PATHWAY THAT PROCEEDS WITHOUT SIGNAL PEPTIDE.
 CC CLEAVAGE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP001118; BAB12796.1;
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR Pfam; PF00006; ATP-synt_ab; 1.
 DR TIGRFAMs; TIGR01026; FLII_yacN; 1.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
 KW Hydrolase; Hydrogen ion transport; ATP synthesis; ATP-binding;
 KW Transport; Protein transport; Flagella; Complete proteome.
 FT NP_BIND 195 202 ATP (BY SIMILARITY).
 SQ SEQUENCE 467 AA; 51948 MW; 0C87A34CDC29A732 CRC64;
 Query Match 96.3%; Score 26; DB 1; Length 467;
 Best Local Similarity 80.0%; Pred. No. 66;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 DB 432 DHAIK 436
 ID GATB BORBU STANDARD; PRT; 485 AA.
 AC O51316.
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glutamyl-tRNA (Gln) amidotransferase subunit B (EC 6.3.5.-) (Glu-ADT
 DE subunit B).
 GN GATB OR B80341.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,

RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi.";
 RL Nature 390:580-586(1997).
 CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED
 CC GLN-tRNA (Gln) THROUGH THE TRANSMUTATION OF MISACLYATED GLU-
 CC tRNA (Gln) IN ORGANISMS WHICH LACK GLUTAMINYL-tRNA SYNTHETASE. THE
 CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH
 CC AN ACTIVATED GAMMA-PHOSPHO-GLU-tRNA (Gln) (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA (Gln) + L-glutamine = ADP
 CC + phosphate + L-glutamyl-tRNA (Gln) + L-glutamate.
 CC -1- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GATB FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE001140; AAC66716.1;
 DR TIGR; BB0341;
 DR InterPro; IPR004413; GatB.
 DR InterPro; IPR003789; GatB_yqey.
 DR InterPro; IPR001773; Gln_amidtransfB.
 DR Pfam; PF01162; GatB; 1.
 DR Pfam; PF02637; DUF186; 1.
 DR Pfam; PF02934; GatB_N; 1.
 DR TIGRFAMs; TIGR00133; gatB; 1.
 DR PROSITE; PS01234; GatB; 1.
 KW Protein biosynthesis; Ligase; Complete proteome.
 SQ SEQUENCE 485 AA; 54520 MW; 622C10C208EDE845 CRC64;
 Query Match 96.3%; Score 26; DB 1; Length 485;
 Best Local Similarity 80.0%; Pred. No. 69;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DHAVK 5
 DB 451 DHAIK 455
 ID FRE6 YEAST STANDARD; PRT; 712 AA.
 AC Q12473.
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ferric reductase transmembrane component 6 precursor (EC 1.6.99.13)
 DE (Ferric-chelate reductase 6).
 GN FRE6 OR YLL051C OR L0593.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Wedler H., Wambutt R.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
 CC -1- COFACTOR: FAD (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE FRE / CYBB FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 247973; CAA89006.1; -
CC EMBL; 273156; CAA97503.1; -
CC SGD; S0003974; FREG.
CC InterPro; IPR002916; Ferric_reduct.
CC Pfam; PF01994; Ferric_reduct; 1.
KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
KW FAD; NAD; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 712 FERRIC REDUCTASE TRANSMEMBRANE COMPONENT
FT 6.
FT NP_BIND 493 499 FAD (POTENTIAL).
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 478 498 POTENTIAL.
FT TRANSMEM 550 570 POTENTIAL.
FT TRANSMEM 553 556 POLY-LEU.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 712 AA; 81989 MW; 5224F12B51544BAA CRC64;

Query Match 96.3%; Score 26; DB 1; Length 712;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:
Db 567 DHAIK 571

RESULT 18
ID CARB_THETN STANDARD; PRT; 1072 AA.
AC Q8BRK0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
GN CARB OR TTE0816.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT Genome Res. 12:689-700(2002).
RL "A complete sequence of T. tengcongensis genome."
CC -1- CATALYTIC ACTIVITY: 2 ADP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -1- COFACTOR: Binds three manganese ions (By similarity).
CC -1- PATHWAY: Arginine biosynthesis.
CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by

the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).
-1- SIMILARITY: BELONGS TO THE CARB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE013048; AAM24073.1; -
CC InterPro; IPR005483; CPase_L1;
CC InterPro; IPR005479; CPase_L_D2.
CC InterPro; IPR005480; CPase_L_D3.
CC InterPro; IPR005481; CPase_L_N.
CC Pfam; PF00289; CPase_L_Chain; 2.
CC Pfam; PF02786; CPase_L_D2; 2.
CC Pfam; PF02787; CPase_L_D3; 1.
CC Pfam; PF02142; MGS; 1.
CC PRINTS; PRO0098; CPASE.
CC PROSITE; PS00866; CPASE 1; 1.
CC PROSITE; PS00867; CPASE 2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 930 1072 ALLOSTERIC DOMAIN.
FT REPEAT 1 546
FT REPEAT 547 1072
FT NP_BIND 153 210 ATP (POTENTIAL).
FT NP_BIND 302 352 ATP (POTENTIAL).
FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1072 AA; 119000 MW; D5F08ADC8BE75D37 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 1072;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
|||:
Db 987 DHAIK 991

RESULT 19
ID YP83 CAEEL STANDARD; PRT; 1244 AA.
AC Q09441;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 139.1 kDa protein C08B11.3 in chromosome II.
GN C08B11.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Simms M.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; Z46676; CAA86663.1; -
 DR WormPep; C08B11.3; CE01473.
 DR InterPro; IPR001606; ARID.
 DR Pfam; PF00096; zf-C2H2; 1.
 DR Pfam; PF01388; ARID; 1.
 DR SMART; SM00501; BRIGHT; 1.
 DR SMART; SM00355; Znf_C2H2; 2.
 DR Hypothetical protein.
 KW Hypothetical protein.
 SQ SEQUENCE 1244 AA; 139117 MW; 50FB3E4C13F73E06 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 1244;
 Best Local Similarity 80.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
 |||:|
 DB 97 DHAIK 101

RESULT 20

ID POLN_EEVV3 STANDARD; PRT; 2485 AA.
 AC P36327;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nonstructural polyprotein (Contains: Nonstructural protein NSP3; Nonstructural protein NSP4).
 DE Nonstructural protein NSP2; Nonstructural protein NSP3; Nonstructural protein NSP4.
 OS Venezuelan equine encephalitis virus (strain 3880).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OK NCBI_TaxID=36382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93079859; PubMed=1448915;
 RA Kinney R.M., Tsuchiya K.R., Snider J.M., Trent D.W.;
 RT "Genetic evidence that epizootic Venezuelan equine encephalitis (VEE) viruses may have evolved from enzootic VEE subtype I-D virus."
 RL Virology 191:569-580(1992).
 CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS BETWEEN THE CODONS FOR 1871-GLN AND 1872-ARG.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; L00930; AAC19324.1; -
 DR EMBL; L00930; AAC19326.1; ALT_SEQ.
 DR PIR; C44213; C44213.
 DR MEROPS; C09.001; -
 DR InterPro; IPR002589; Alpp.
 DR InterPro; IPR002620; Peptidase C9.
 DR InterPro; IPR001788; RNA_dep_RNAPol2.
 DR InterPro; IPR000606; Viral_helicase1.
 DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR Pfam; PF01661; Alpp; 1.
 DR Pfam; PF01707; Peptidase C9; 1.
 DR SMART; SM00506; Alpp; 1.
 KW Polyprotein; Nonstructural protein; RNA-binding; ATP-binding; Helicase.

FT CHAIN 1 535 NONSTRUCTURAL PROTEIN NSP1.
 FT CHAIN 536 1329 NONSTRUCTURAL PROTEIN NSP2.
 FT CHAIN 1330 1871 NONSTRUCTURAL PROTEIN NSP3.
 FT CHAIN 1872 2485 NONSTRUCTURAL PROTEIN NSP4.
 FT NP BIND 721 728 ATP (POTENTIAL).
 SQ SEQUENCE 2485 AA; 277282 MW; 83458A9646D6D639 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 2485;
 Best Local Similarity 80.0%; Pred. No. 4.3e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
 |||:|
 DB 1230 DHAIK 1234

RESULT 21

ID POLN_EEVVP STANDARD; PRT; 2492 AA.
 AC P36328;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nonstructural polyprotein (Contains: Nonstructural protein NSP1; Nonstructural protein NSP2; Nonstructural protein NSP3; Nonstructural protein NSP4).
 DE Venezuelan equine encephalitis virus (strain P676).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OK NCBI_TaxID=36385;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93079859; PubMed=1448915;
 RA Kinney R.M., Tsuchiya K.R., Snider J.M., Trent D.W.;
 RT "Genetic evidence that epizootic Venezuelan equine encephalitis (VEE) viruses may have evolved from enzootic VEE subtype I-D virus."
 RL Virology 191:569-580(1992).
 CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS BETWEEN THE CODONS FOR 1879-GLN AND 1880-ARG.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; L04653; CAB23722.1; -
 DR PIR; A44213; A44213.
 DR MEROPS; C09.001; -
 DR InterPro; IPR002589; Alpp.
 DR InterPro; IPR002620; Peptidase C9.
 DR InterPro; IPR001788; RNA_dep_RNAPol2.
 DR InterPro; IPR000606; Viral_helicase1.
 DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR Pfam; PF01661; Alpp; 1.
 DR Pfam; PF01707; Peptidase C9; 1.
 DR SMART; SM00506; Alpp; 1.
 KW Polyprotein; Nonstructural protein; RNA-binding; ATP-binding; Helicase.

FT CHAIN 1 535 NONSTRUCTURAL PROTEIN NSP1.
 FT CHAIN 536 1329 NONSTRUCTURAL PROTEIN NSP2.
 FT CHAIN 1330 1879 NONSTRUCTURAL PROTEIN NSP3.
 FT CHAIN 1880 2492 NONSTRUCTURAL PROTEIN NSP4.
 FT NP BIND 721 728 ATP (POTENTIAL).
 SQ SEQUENCE 2492 AA; 277836 MW; E55D056CD7AED6EE CRC64;

Query Match 96.3%; Score 26; DB 1; Length 2492;
 Best Local Similarity 80.0%; Pred. No. 4.3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

Db 1230 DHAIK 1234

RESULT 22

POLN EVVVT STANDARD; PRT; 2492 AA.
AC P27282; 1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nonstructural polyprotein [Contains: Nonstructural protein NSP1;
DE Nonstructural protein NSP2; Nonstructural protein NSP3; Nonstructural
DE protein NSP4].
OS Venezuelan equine encephalitis virus (strain Trinidad donkey).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11038;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89243175; PubMed=2524126;
RA Kinney R.M., Johnson B.J.B., Welch J.B., Tsuchiya K.R., Trent D.W.;
RT "The full-length nucleotide sequences of the virulent Trinidad donkey
RT strain of Venezuelan equine encephalitis virus and its attenuated
RT vaccine derivative, strain TC-83";
RL Virology 170:19-30(1989).
CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
CC BETWEEN THE CODONS FOR 1879-GLN AND 1880-ARG.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J04332; AAB02518.1; -
CC PIR; A31467; MNWVTD.
CC MEROPS; C09.001; -
CC InterPro; IPR002589; Alpp.
CC InterPro; IPR002620; Peptidase_C9.
CC InterPro; IPR001788; RNA_dep_RNAPol2.
CC InterPro; IPR000606; Viral_helicase1.
CC Pfam; PF00978; RNA_dep_RNAPol2; 1.
CC Pfam; PF01443; Viral_helicase1; 1.
CC Pfam; PF01661; Alpp; 1.
CC Pfam; PF01707; Peptidase_C9; 1.
CC SMART; SM00506; Alpp; 1.
KW Polyprotein; Nonstructural protein; RNA-binding; ATP-binding;
KW Helicase.
FT CHAIN 1 535 NONSTRUCTURAL PROTEIN NSP1.
FT CHAIN 536 1329 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 1330 1879 NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 1880 2492 NONSTRUCTURAL PROTEIN NSP4.
FT NP BIND 721 728 ATP (POTENTIAL).
SQ SEQUENCE 2492 AA; 277902 MW; 1BAD415B70DC3FA0 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 2492;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

Db 1230 DHAIK 1234

RESULT 23

GLNB RHET

ID GLNB RHET STANDARD; PRT; 112 AA.

AC O54053;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nitrogen regulatory protein P-II.

GN GLNB.

OS Rhizobium etli.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=29449;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CE3;

RA Tate R., Riccio A., Merrick M., Patriarca E.J.;

RT "The Rhizobium etli amtB gene coding for an NH4+ transporter is

RT down-regulated early during bacteroid differentiation";

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN

CC TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP. P-

CC II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS).

CC THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II

CC IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS (BY

CC SIMILARITY).

CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS

CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO

CC NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II

CC IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED (BY

CC SIMILARITY).

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE P(II) PROTEIN FAMILY.

CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ002489; CAA05496.1; ALT_INIT.
CC HSSP; P38504; 1GNK.
CC InterPro; IPR002332; PII_GlnB_UMP.
CC InterPro; IPR002187; PII_glnB.
CC Pfam; PF00543; P-II; 1.
CC PRINTS; PR00340; PIIGLNB.
CC ProDom; PD001194; PII_GlnB; 1.
CC PROSITE; PS00496; PII_GlnB_UMP; FALSE_NEG.
CC PROSITE; PS00638; PII_GlnB_CTER; 1.
KW Transcription regulation; Nitrogen fixation.
FT BINDING 51 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12134 MW; DFB794E29683E0 CRC64;

Query Match 88.9%; Score 24; DB 1; Length 112;
Best Local Similarity 80.0%; Pred. No. 41;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

Db 97 DHAIV 101

RESULT 24

YCBM_BACSU

ID YCBM_BACSU STANDARD; PRT; 158 AA.

AC P42245;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein ycbm.

GN YCBM.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

RN NCBI_TaxID=1423;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst E., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borriess R., Bourlier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Duisterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,

RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Kassarjian A., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maubel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,

RA Rieger M., Rivolta C., Roche E., Roche M., Rose M., Sadaie Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weltenegeger T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

RT subtilis".

RL Nature 390:249-256(1997).

CC -!- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC EMBL; D30808; BAA06477.1; -

DR EMBL; D30808; BAA06477.1; -

DR Subtilisin; EG1118; ycdM.

DR InterPro; IPR003594; ATPbind_ATPase.

DR InterPro; IPR004359; HIS_KIN_sig.

DR Pfam; PF02518; HATPase_C; 1.

DR SMART; SM00387; HATPase_C; 1.

DR PROSITE; PS50109; HIS_KIN; 1.

DR Hypothetical protein; Complete proteome.

FW DOMAIN 1 157 HISTIDINE KINASE.

FT SEQUENCE 158 AA; 17994 MW; E605F5702F6F5C40 CRC64;

SQ

Query Match 88.9%; Score 24; DB 1; Length 158;

Best Local Similarity 80.0%; Pred. No. 60;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DHAVK 5

||||

Db 55 DHALK 59

RESULT 25

RPSH_PSEAE

ID RPSH_PSEAE STANDARD; PRT; 193 AA.

AC Q06198;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE RNA polymerase sigma-H factor (Sigma-30).

GN ALGU OR ALGT OR PA0762.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93163045; PubMed=8432708;

RA Martin D.W., Holloway B.W., Deretic V.;

RA "Characterization of a locus determining the mucoid status of

RA Pseudomonas aeruginosa: ALGU shows sequence similarities with a

RA Bacillus sigma factor."

RL J. Bacteriol. 175:1153-1164(1993).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=PAO381, and PAO568;

RA MEDLINE=93391358; PubMed=8378309;

RA Martin D.W., Schurr M., Mudd M.H., Govan J.R., Holloway B.W.,

RA Deretic V.;

RA "Mechanism of conversion to mucoidy in Pseudomonas aeruginosa

RA infecting cystic fibrosis patients."

RL Proc. Natl. Acad. Sci. U.S.A. 90:8377-8381(1993).

RN [3]

RP SEQUENCE FROM N.A.

RX STRAIN=FRD1;

RA MEDLINE=95050229; PubMed=7961421;

RA Devries C.A., Oman D.E.;

RA "Mucoid-to-nonmucoid conversion in alginate-producing Pseudomonas

RA aeruginosa often results from spontaneous mutations in algT, encoding

RA a putative alternate sigma factor, and shows evidence for

RA autoregulation."

RL J. Bacteriol. 176:6677-6687(1994).

RN [4]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 15692 / PA01;

RA MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman K., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen."

RL Nature 406:959-964(2000).

CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES

CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND

CC THEN IS RELEASED. THIS SIGMA FACTOR REGULATES GENES SUCH AS ALGU,

CC INVOLVED IN ALGINATE BIOSYNTHESIS.

CC -!- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY. ECF SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC EMBL; L14760; AAA87628.1; -

DR EMBL; L14760; AAA87628.1; -

DR EMBL; L14761; AAA87631.1; -

DR EMBL; U49151; AAC43714.1; -

DR EMBL; L36379; AAA64439.1; --
 DR EMBL; AE004511; AAG04151.1; --
 DR PIR; A49704; A49704; Sigma70_ECF.
 DR InterPro; IPR000838; Sigma70_ECF.
 DR Pfam; PF00776; Sigma70_ECF; 1.
 DR PROSITE; PS01063; SIGMA70_ECF; 1.
 KW Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
 KW DNA-binding; Complete proteome.
 FT DOMAIN 49 62 POLYMERASE CORE BINDING (POTENTIAL).
 FT DNA_BIND 157 176 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 193 AA; 22196 MW; 9C10419C51F29C64 CRC64;

Query Match 88.9%; Score 24; DB 1; Length 193;
 Best Local Similarity 80.0%; Pred. No. 75;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
 DB 110 DHALK 114

Search completed: June 20, 2003, 20:52:57
 Job time : 8.02273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:42:49 ; Search time 26.3636 Seconds
(without alignments)
39.078 Million cell updates/sec

Title: US-10-105-008-14

Perfect score: 27

Sequence: 1 DHAVK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

SPTREMBL.21.*
1: sp archaea.*
2: sp bacteria.*
3: sp fungi.*
4: sp human.*
5: sp invertebrate.*
6: sp mammal.*
7: sp mhc.*
8: sp organelle.*
9: sp phase.*
10: sp plant.*
11: sp rodent.*
12: sp virus.*
13: sp vertebrate.*
14: sp unclassified.*
15: sp virus.*
16: sp bacterioph.*
17: sp archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	101	2 Q8VVM9	Q8VVM9 vibrio chol
2	27	100.0	155	10 Q64971	Q64971 arabidopsis
3	27	100.0	161	10 Q9LEP5	Q9LEP5 brassica ol
4	27	100.0	190	10 Q9SPG4	Q9SPG4 arabidopsis
5	27	100.0	199	10 Q39270	Q39270 brassica ca
6	27	100.0	205	10 Q39268	Q39268 brassica ca
7	27	100.0	206	5 Q19916	Q19916 caenorhabdi
8	27	100.0	207	5 Q8VYX5	Q8VYX5 brassica ol
9	27	100.0	250	10 Q9LJ08	Q9LJ08 oryza sativ
10	27	100.0	250	16 Q822T4	Q822T4 salmonella
11	27	100.0	263	2 Q9ADT4	Q9ADT4 salmonella
12	27	100.0	263	16 Q986V1	Q986V1 rhizobium 1
13	27	100.0	280	2 Q9L7R7	Q9L7R7 salmonella
14	27	100.0	285	16 Q82KT8	Q82KT8 salmonella
15	27	100.0	315	2 Q9X588	Q9X588 neisseria f
16	27	100.0	317	17 Q8TNE2	Q8TNE2 methanosarc

RESULT 1

Q8VVM9 PRELIMINARY; PRT; 101 AA.
AC Q8VVM9;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 11.1 kDa protein.
OS Vibrio cholerae.

ALIGNMENTS

17	27	100.0	327	2 Q8VW39	Q8VW39 klebsiella
18	27	100.0	334	17 Q8TLD2	Q8TLD2 methanosarc
19	27	100.0	351	17 Q980F5	Q980F5 sulfolobus
20	27	100.0	407	3 Q95948	Q95948 saccharomyc
21	27	100.0	434	11 Q9D3W6	Q9D3W6 mus musculu
22	27	100.0	459	2 P94171	P94171 arcanobacte
23	27	100.0	495	16 Q99X54	Q99X54 staphylococ
24	27	100.0	502	3 Q8X0F5	Q8X0F5 neurospora
25	27	100.0	507	16 Q8Y287	Q8Y287 ralstonia s
26	27	100.0	508	15 Q8X718	Q8X718 escherichia
27	27	100.0	552	12 Q91CM0	Q91CM0 antheraea p
28	27	100.0	553	12 Q9WGG2	Q9WGG2 hyphantria
29	27	100.0	581	17 Q29674	Q29674 archaeoglob
30	27	100.0	629	10 Q9LXM7	Q9LXM7 arabidopsis
31	27	100.0	631	10 Q8W4D0	Q8W4D0 arabidopsis
32	27	100.0	632	15 Q8U7R3	Q8U7R3 agrobacteri
33	27	100.0	723	16 Q9KVP6	Q9KVP6 vibrio chol
34	27	100.0	798	3 P79020	P79020 emericella
35	27	100.0	818	3 Q43059	Q43059 schizosacch
36	27	100.0	885	5 Q9VK27	Q9VK27 drosophila
37	27	100.0	1130	10 Q9FIV9	Q9FIV9 arabidopsis
38	27	100.0	1245	12 Q8QTI5	Q8QTI5 sindbis vir
39	27	100.0	1531	2 Q9RNZ7	Q9RNZ7 zymomonas m
40	27	100.0	1545	2 Q9RNL3	Q9RNL3 zymomonas m
41	27	100.0	2248	5 Q9VY17	Q9VY17 drosophila
42	26	96.3	110	2 Q8VPA1	Q8VPA1 enterococcu
43	26	96.3	134	12 Q9WR11	Q9WR11 west africa
44	26	96.3	167	2 Q9RNF8	Q9RNF8 helicobacte
45	26	96.3	167	2 Q89434	Q89434 helicobacte
46	26	96.3	167	2 Q86247	Q86247 helicobacte
47	26	96.3	167	2 Q8VP42	Q8VP42 helicobacte
48	26	96.3	180	2 Q32593	Q32593 enterococcu
49	26	96.3	202	10 Q9FWQ7	Q9FWQ7 arabidopsis
50	26	96.3	227	10 Q9SKK9	Q9SKK9 arabidopsis
51	26	96.3	292	2 Q9FDS9	Q9FDS9 streptococc
52	26	96.3	297	16 Q9A098	Q9A098 streptococc
53	26	96.3	348	17 Q971H7	Q971H7 sulfolobus
54	26	96.3	400	8 Q9MJ78	Q9MJ78 physarum po
55	26	96.3	442	16 Q8R6B4	Q8R6B4 fusobacteri
56	26	96.3	465	17 Q9U2B9	Q9U2B9 pyrococcus
57	26	96.3	509	4 Q9NUJ3	Q9NUJ3 homo sapien
58	26	96.3	525	16 Q99RH3	Q99RH3 staphylococ
59	26	96.3	533	6 Q95JY2	Q95JY2 macaca fasc
60	26	96.3	597	10 Q9LQ68	Q9LQ68 arabidopsis
61	26	96.3	599	17 Q9HK38	Q9HK38 thermoplas
62	26	96.3	700	4 Q8TC20	Q8TC20 homo sapien
63	26	96.3	750	5 Q9VUD5	Q9VUD5 drosophila
64	26	96.3	840	6 Q95JRO	Q95JRO macaca fasc
65	26	96.3	915	5 Q9VDK3	Q9VDK3 drosophila
66	26	96.3	963	3 Q9C2Y9	Q9C2Y9 emericella
67	26	96.3	998	17 Q8THF6	Q8THF6 methanosarc
68	26	96.3	1072	16 Q8R8K0	Q8R8K0 thermosuaer
69	26	96.3	1394	5 Q9VS89	Q9VS89 drosophila
70	26	96.3	1879	12 Q9WJ31	Q9WJ31 venezuelan
71	26	96.3	1879	12 Q8QKW7	Q8QKW7 venezuelan
72	26	96.3	2447	12 Q9WJD0	Q9WJD0 venezuelan
73	26	96.3	2451	12 Q9WJD4	Q9WJD4 venezuelan
74	26	96.3	2455	12 Q9WJC8	Q9WJC8 venezuelan
75	26	96.3	2477	12 Q9WJD1	Q9WJD1 venezuelan

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=MO10; TRANSPOSON=SXT ELEMENT;
RX MEDLINE=21485931; PubMed=11600347;
RA Hochhut B., Lotfi Y., Mazel D., Faruque S.M., Woodgate R.,
RA Waldor M.K.;
RT Molecular Analysis of Antibiotic Resistance Gene Clusters in Vibrio
RT cholerae O139 and O1 SXT Constains.;
RL Antimicrob. Agents Chemother. 45:2991-3000(2001).
DR EMBL; AY034138; AAK64588.1; -;
DR InterPro; IPR000485; ASNC trans_reg.
DR Pfam; PF00126; HTH_1; 1.
DR PRINTS; PR00033; HTHASNC.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 101 AA; 11146 MW; 2FC2DCF191BDE4E8 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 94 DHAVK 98

RESULT 2
OC 064971
ID AC 064971 PRELIMINARY; PRT; 155 AA.
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative disease resistance protein (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=98191999; PubMed=9530866;
RA Aarts M.G., te Lintel Hekkert B., Holub E.B., Beynon J.L.,
RA Stiekema W.J., Pereira A.;
RT Identification of R-gene homologous DNA fragments genetically linked
RT to disease resistance loci in Arabidopsis thaliana.;
RL Mol. Plant Microbe Interact. 11:251-258(1998).
DR EMBL; AF039377; AAC14553.1; -;
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
FT NON_TER 155
FT NON_TER 155
SQ SEQUENCE 155 AA; 17999 MW; 0F40F3BF63A9395B CRC64;

Query Match 100.0%; Score 27; DB 10; Length 155;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 55 DHAVK 59

RESULT 3
OC 09LEP5
ID AC 09LEP5 PRELIMINARY; PRT; 161 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-JUN-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Hypothetical 18.5 kDa protein (Fragment).
GN RGL-LD.
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RA Vicente J.G., King G.J.;
RT Characterisation of disease resistance gene-like sequences in
RT Brassica oleracea L.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250322; CAB94290.1; -;
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
KW Hypothetical protein.
FT NON_TER 161
FT NON_TER 161
SQ SEQUENCE 161 AA; 18525 MW; 63C09D24D4B25B4B CRC64;

Query Match 100.0%; Score 27; DB 10; Length 161;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 56 DHAVK 60

RESULT 4
OC 09SPG4
ID AC 09SPG4 PRELIMINARY; PRT; 190 AA.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative transcription factor.
GN MYB89.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=99056848; PubMed=9839469;
RA Kranz H.D., Denekamp M., Greco R., Jin H., Leyva A., Meissner R.C.,
RA Petroni K., Urzaizqui A., Bevan M., Martin C., Smeekens S.,
RA Tonelli C., Paz-Ares J., Weisshaar B.;
RT Towards functional characterisation of the members of the R2R3-MYB
RT gene family from Arabidopsis thaliana.;
RL Plant J. 16:263-276(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Stracke R., Weisshaar B.;
RT "R2R3-MYB transcription factor gene nomenclature in Arabidopsis
RT thaliana.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneo T., Miyajima N.,
RA Tabata S.;
RT Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned pl and TAC clones.;
RT DNA Res. 5:203-216(1998).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR ENBL; ARL75995; AAD53100.1; -;
DR ENBL; AB012243; BAB08902.1; -;
DR HSP; P01103; IPOW.
DR InterPro; IPR001005; MYB_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00395; SANT; 2.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
DR PROSITE; PS00334; MYB_2; UNKNOWN_2.
DR PROSITE; PS00090; MYB_3; 2.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 190 AA; 22676 MW; 64988B1D526FF315 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

Db 142 DHAVK 146

RESULT 5

ID Q39270 PRELIMINARY; PRT; 199 AA.
AC Q39270;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cysteine proteinase inhibitor.
OS Brassica campestris (field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96178877; PubMed=8616262;
RA Lim C.O., Lee S.I., Chung W.S., Park S.H., Hwang I., Cho M.J.;
RT "Characterization of a cDNA encoding cysteine proteinase inhibitor
from Chinese cabbage (Brassica campestris L. ssp. pekinensis) flower
buds.";
RL Plant Mol. Biol. 30:373-379(1996).
DR ENBL; L41355; AAC37479.1; -;
DR HSP; P01038; ICEW.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; cystatin; 2.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 2.
DR PROSITE; PS00287; CYSTATIN; UNKNOWN_1.
SQ SEQUENCE 199 AA; 22519 MW; 5B1C0E2847A9ED6D CRC64;

Query Match 100.0%; Score 27; DB 10; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

Db 126 DHAVK 130

RESULT 6

ID Q39268 PRELIMINARY; PRT; 205 AA.
AC Q39268;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cysteine proteinase inhibitor.
GN BCPI-2.
OS Brassica campestris (field mustard).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RA Lim C.O., Beachy R.N., Cho M.J.;
RT "Cysteine proteinase inhibitor cDNAs from Chinese cabbage.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR ENBL; U51119; AAA96316.1; -;
DR HSP; P01038; ICEW.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; cystatin; 2.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 2.
DR PROSITE; PS00287; CYSTATIN; UNKNOWN_1.
SQ SEQUENCE 205 AA; 22945 MW; 089D1CE748A8F685 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

Db 132 DHAVK 136

RESULT 7

ID Q19916 PRELIMINARY; PRT; 206 AA.
AC Q19916;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 24.3 kDa protein.
GN F31A3.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA Murray J., Lee T.T.;
RT "The sequence of C. elegans cosmid F31A3.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; U58742; AAB36857.1; -;
KW Hypothetical protein.
SQ SEQUENCE 206 AA; 24293 MW; 2A8835FA618388A8 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5

Db 179 DHAVK 183

RESULT 8

QSVYX5
ID Q8VYX5 PRELIMINARY; PRT; 207 AA.
AC Q8VYX5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cysteine protease inhibitor CPI-1.
GN CPI-1.
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RA Watson L.M., Coupe S.A.;
RT "Cloning and characterization of a cysteine protease inhibitor from
RT senescing broccoli florets."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY065838; AAL59842.1; -
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 2.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 2.
DR PROSITE; PS00287; CYSTATIN; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 207 AA; 23317 MW; DID77FF9A03BB0B7 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5

Db 135 DHAVK 139

RESULT 9

Q9LJ08
ID Q9LJ08 PRELIMINARY; PRT; 250 AA.
AC Q9LJ08;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE ESTs AU067919 (C10906).
OS Oryza sativa (Rice)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0667A10."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001073; BAA89582.1; -
DR ANU-2DPAGE; Q9LJ08; -
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 2.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; UNKNOWN_1.
SQ SEQUENCE 250 AA; 27269 MW; 6F2A7D4EB8BAF66F CRC64;

Query Match 100.0%; Score 27; DB 10; Length 250;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5

Db 173 DHAVK 177

RESULT 10

Q8Z2T4
ID Q8Z2T4 PRELIMINARY; PRT; 250 AA.
AC Q8Z2T4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein STY3858.
GN STY3858.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dow L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL627279; CAD09606.1; -
DR InterPro; IPR001823; Aldol_epimerase.
DR Pfam; PF01283; Aldose_epim; 1.
DR PROSITE; PS00287; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 250 AA; 27456 MW; E9F943225CDA9EB3 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 250;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5

Db 202 DHAVK 206

RESULT 11

Q9ADT4
ID Q9ADT4 PRELIMINARY; PRT; 263 AA.
AC Q9ADT4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative LysR-type transcriptional regulator.
GN QACEDELTA1.
OS Salmonella enterica subsp. enterica serovar Typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=90371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96-5227;
RA MEDLINE-21429247; PubMed=11544236;
RA Boyd D., Peters G.A., Cloeckaert A., Boumedine K.S.,
RA Chaslus-Dancla E., Imberechts H., Mulvey M.R.;
RT "Complete Nucleotide Sequence of a 43-Kilobase Genomic Island
RT Associated with the Multidrug Resistance Region of Salmonella enterica
RT Serovar Typhimurium DT104 and Its Identification in Phage Type DT120
RT and Serovar Agona";
RL J. Bacteriol. 183:5725-5732(2001).

-1- SIMILARITY: TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.

```
CC EMBL; AF261825; AAK02052.1; -
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN 1.
DR DNA-binding; Transcription regulation.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 263 AA; 28647 MW; 69E3CD4B1F190062 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 97 DHAVK 101

RESULT 12
Q986V1 PRELIMINARY; PRT; 263 AA.
AC Q986V1
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE 3-alpha-hydroxysteroid dehydrogenase.
GN MLF17199.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OC NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
Takeuchi C., Yanada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003011; BAB53352.1; -
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 263 AA; 27057 MW; CD5349E3A69D231D CRC64;

Query Match 100.0%; Score 27; DB 16; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 149 DHAVK 153

RESULT 13
Q9L7R7 PRELIMINARY; PRT; 280 AA.
AC Q9L7R7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative aldose-1-epimerase (EC 5.1.3.3).
GN YHR.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
```

```
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RA Krogan N.J., Zhang R., Neuhard J., Kelln R.A.;
RT "Utilization of Dihydroorotate as Sole Pyrimidine Source by Salmonella
typhimurium.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF220438; AAF27924.1; -
DR InterPro; IPR001823; Aldi-epimerase.
DR Pfam; PF01263; Aldose_epim; 1.
KW Isomerase.
SQ SEQUENCE 280 AA; 30511 MW; C631D6793D914DBB CRC64;

Query Match 100.0%; Score 27; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 197 DHAVK 201

RESULT 14
Q8ZKT8 PRELIMINARY; PRT; 285 AA.
AC Q8ZKT8
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative aldose-1-epimerase.
GN YHR OR STM4020.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.R., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AB008887; AAL22859.1; -
DR InterPro; IPR001823; Aldi-epimerase.
DR Pfam; PF01263; Aldose_epim; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 285 AA; 31071 MW; DA96EF2B3698612C CRC64;

Query Match 100.0%; Score 27; DB 16; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 202 DHAVK 206

RESULT 15
Q9X588 PRELIMINARY; PRT; 315 AA.
AC Q9X588
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Porin (Fragment).
GN POR.
```

OS Neisseria flava.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=34026;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC14221;
 RX MEDLINE=99242827; PubMed=10225902;
 RA Derrick J.P., Urwin R., Suker J., Feavers I.M., Maiden M.C.;
 RT "Structural and evolutionary inference from molecular variation in
 Neisseria porins.";
 RL Infect. Immun. 67;2406-2413(1999).
 DR EMBL; AF121873; AAD29370.1; -;
 DR InterPro; IPR001702; Porin gram-ve.
 DR Pfam; PF00267; Gram-ve porins; 1.
 DR PRINTS; PR00182; ECOLNIPORIN.
 DR NON_TER 1
 FT NON_TER 315 315
 SQ SEQUENCE 315 AA; 34245 MW; FFB46A457341F6CD CRC64;

Query Match 100.0%; Score 27; DB 2; Length 315;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DHAVK 5
 Db 145 DHAVK 149

RESULT 16

Q8TNE2 PRELIMINARY; PRT; 317 AA.
 AC Q8TNE2;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein MA2346.
 GN MA2346.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nussbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity.";
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE010923; AAM05736.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 317 AA; 35297 MW; 565B274BFE418441 CRC64;

Query Match 100.0%; Score 27; DB 17; Length 317;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DHAVK 5
 Db 210 DHAVK 214

RESULT 17

Q8VW39 PRELIMINARY; PRT; 327 AA.
 AC Q8VW39;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative regulator protein.
 OS Klebsiella pneumoniae.
 OC Plasmid R55
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21344636; PubMed=11451703;
 RA Cloeckaert A., Baucheron S., Chaslus-Dancla E.;
 RT "Nonenzymatic Chloramphenicol Resistance Mediated by IncC Plasmid R55
 is Encoded by a *flor* Gene Variant.";
 RL Antimicrob. Agents Chemother. 45:2381-2382(2001).
 DR EMBL; AF332662; AAL33887.1; -;
 DR InterPro; IPR000847; HTH_Lysr.
 DR InterPro; IPR005119; Lysr_subst.
 DR Pfam; PF00126; HTH_1; 1.
 DR Pfam; PF03466; Lysr substrate; 1.
 DR PRINTS; PR00039; HTHLYSR.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
 KW Plasmid.
 SQ SEQUENCE 327 AA; 36031 MW; C29477AA88240423 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 327;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
 Db 94 DHAVK 98

RESULT 18

Q8TLD2 PRELIMINARY; PRT; 334 AA.
 AC Q8TLD2;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein MA3104.
 GN MA3104.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nussbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity.";
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE010104; AAM06477.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 334 AA; 37474 MW; DEA9632B8A0EAL7B CRC64;

Query Match 100.0%; Score 27; DB 17; Length 334;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
 Db 227 DHAVK 231

RESULT 19

Q980F5 PRELIMINARY; PRT; 351 AA.
 AC Q980F5;
 DT 01-OCT-2001 (TREMELrel. 18, Created)
 DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE Phosphate regulatory protein, putative.
 GN SS00356.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=2133296; PubMed=11427726;
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Awaysz M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
 DR ENBL; AE006669; AAK40687.1;
 DR InterPro; IPR002754; Phou.
 DR ProDom; PD014067; DUF65; 1.
 KW Complete proteome.
 SQ SEQUENCE 351 AA; 40385 MW; A0A8316697E1EA2F CRC64;

Query Match 100.0%; Score 27; DB 17; Length 351;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
 Db 264 DHAVK 268

RESULT 20

Q05948 PRELIMINARY; PRT; 407 AA.
 AC Q05948;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE Chromosome XII COSMID 8033.
 GN YLR225C OR L8083.10.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=97313267; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,
 RA Benes V., Bruckner M., Dalius H., Dubois E., Dusterhoft A.,
 RA Entian K.D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Heuss-Neitzel D., Hilbert F., Hilger F., Kleine K., Kotter P.,
 RA Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
 RA Muller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,

RA Fortetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urrestarazu L.A., Vandenberg M., Verhasselt P.,
 RA Vierdeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
 RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Honeisel J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RL Nature 387:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RC Kallworth K.;
 RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Waterston R.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Cherry J.M.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR ENBL; U19027; AAB67413.1;
 DR SGD; S0004215; YLR225C.
 SQ SEQUENCE 407 AA; 46423 MW; B251F1AAE421DBDA CRC64;

Query Match 100.0%; Score 27; DB 3; Length 407;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DHAVK 5
 Db 320 DHAVK 324

RESULT 21

Q9D3W6 PRELIMINARY; PRT; 434 AA.
 AC Q9D3W6;
 DT 01-JUN-2001 (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE 4933430F16RIK protein.
 GN 4933430F16RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Rodriguez I., Sakamoto N.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).

```

DR EMBL; AK016991; BAB30542.1; -
DR MGD; MGI:1918511; 4933430F16Pik.
DR InterPro; IPR000408; Reg_chx_condens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PD00320; GPROTEINBRPT.
DR PRODOM; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS00626; RCC1.2; UNKNOWN_1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS0082; WD_REPEATS_2; 6.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
DR Repeat; WD repeat.
KW REPEAT; WD repeat.
SQ SEQUENCE 434 AA; 48965 MW; 7E35C651895F318E CRC64;

Query Match 100.0%; Score 27; DB 11; Length 434;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 165 DHAVK 169

RESULT 22
P94171 PRELIMINARY; PRT; 459 AA.
AC P94171;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Rep.
GN REP.
OS Arcanobacterium pyogenes.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Actinomycineae; Actinomycetaceae; Arcanobacterium.
OX NCBI_TaxID=1661;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BBR1;
RX MEDLINE=98292760; PubMed=9620977;
RA Billington S.J., Jost B.H., Songer J.G.;
RT "The Arcanobacterium (Actinomycetes) pyogenes plasmid pAP1 is a member
of the pU101/pU11 family of rolling circle replication plasmids.";
RL J. Bacteriol. 180:3233-3236(1998).
DR EMBL; U83788; AAC46399.1; -
DR InterPro; IPR000989; Rep.
DR Pfam; PF01446; Rep; 1.
KW Plasmid.
SQ SEQUENCE 459 AA; 51264 MW; 2A0A4B0A043BF865 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 169 DHAVK 173

RESULT 23
Q99X54 PRELIMINARY; PRT; 495 AA.
AC Q99X54;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE AldA protein (Aldehyde dehydrogenase homologue).
GN ALDA OR SAV0167 OR SA0162.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

```

```

OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311352; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sakimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hirumatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003358; BAB56329.1; -
DR EMBL; AP003129; BAB41382.1; -
DR HSSP; P05091; ICW3.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 495 AA; 53659 MW; 36B448F85BFECDD2 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 53 DHAVK 57

RESULT 24
Q8X0F5 PRELIMINARY; PRT; 502 AA.
AC Q8X0F5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 50.3 kDa protein.
GN B14A6.010.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Partmann B., Holland R.,
Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL670007; CAD21303.1; -
KW Hypothetical protein.
SQ SEQUENCE 502 AA; 50255 MW; A818614A080B2CAD CRC64;

Query Match 100.0%; Score 27; DB 3; Length 502;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DHAVK 5
Db 79 DHAVK 83

RESULT 25
Q8Y287 PRELIMINARY; PRT; 507 AA.
ID Q8Y287

```


Q8Y287;
 AC 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Probable threonine dehydratase (Threonine deaminase) transmembrane
 DE protein (EC 4.2.1.16).
 GN ILVA OR RSC0449 OR RS04453.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OC NCBL_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMI1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Sigquier P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 ET "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR EMBL; AL646059; CAD13977.1; -.
 DR InterPro; IPR001926; B6 enzyme beta.
 DR InterPro; IPR000634; S/T dehydratase.
 DR InterPro; IPR001721; ThrDh_C.
 DR Pfam; PF00291; PALP; 1.
 DR Pfam; PF00585; Thr dehydrat C; 2.
 DR TIGRFAMs; TIGR01124; ilva_2Cterm; 1.
 DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
 KW Lyase; Complete proteome.
 SQ SEQUENCE 507 AA; 55666 MW; 1C7A04667BC5F644 CRC64;

Query Match 100.0%; Score 27; DB 16; Length 507;
 Best Local Similarity 100.0%; Pred. NO. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRAVK 5
 Db 132 DRAVK 136

Search completed: June 20, 2003, 20:57:00
 Job time : 30.3636 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 31.9318 Seconds

(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-16

Perfect score: 26

Sequence: 1 KHAVE 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	5	20	AAV17156
2	26	100.0	5	20	AAW91035
3	26	100.0	5	21	AAW27348
4	26	100.0	5	21	AAV73805
5	26	100.0	5	22	AAW65376
6	26	100.0	5	23	AAW47483
7	26	100.0	31	22	ABG01734
8	26	100.0	213	22	ABG01735
9	26	100.0	266	22	ABA71701
10	26	100.0	332	22	AAW99734

11	26	100.0	366	21	AAV54091
12	26	100.0	366	21	AAV43793
13	26	100.0	377	23	AAE21827
14	26	100.0	378	23	AAE21814
15	26	100.0	454	21	AAV74537
16	26	100.0	454	21	AAV74538
17	26	100.0	454	21	AAV74539
18	26	100.0	461	21	AAV25591
19	26	100.0	461	21	AAV25515
20	26	100.0	461	23	AAE20317
21	26	100.0	461	23	AAE20321
22	26	100.0	466	21	AAV99413
23	26	100.0	466	22	AAU29181
24	26	100.0	466	22	AAV74731
25	26	100.0	466	22	AAV87577
26	26	100.0	466	22	AAE66162
27	26	100.0	497	23	AAU79171
28	26	100.0	500	21	AAV99385
29	26	100.0	500	22	AAV87567
30	26	100.0	500	22	AAE66134
31	26	100.0	500	23	AAE95523
32	26	100.0	500	23	AAE84917
33	26	100.0	500	23	AAU83681
34	26	100.0	588	20	AAV17297
35	26	100.0	588	21	AAV82219
36	26	100.0	1039	22	AAE69116
37	26	100.0	1040	21	AAV93291
38	26	100.0	1047	22	AAE10021
39	26	100.0	1048	13	AAV20637
40	26	100.0	1048	18	AAV93268
41	26	100.0	1067	21	AAV93268
42	26	100.0	1082	22	AAU27568
43	26	100.0	1082	22	AAE10020
44	26	100.0	1312	22	AAU27576
45	26	100.0	1312	22	AAU27581
46	26	100.0	1312	22	AAE10023
47	26	100.0	1312	22	AAE10030
48	26	100.0	1389	22	AAU27605
49	26	100.0	1389	22	AAE10040
50	26	100.0	1392	22	AAU27578
51	26	100.0	1392	22	AAE10025
52	26	100.0	1403	22	AAU27608
53	26	100.0	1403	22	AAE10043
54	26	100.0	1443	22	AAE68472
55	26	100.0	1444	22	AAU27602
56	26	100.0	1444	22	AAE10037
57	26	100.0	1447	22	AAU27577
58	26	100.0	1447	22	AAE10024
59	26	100.0	1474	22	AAU27575
60	26	100.0	1474	22	AAE10022
61	26	100.0	1666	22	AAE59298
62	26	100.0	1717	22	AAE20498
63	26	100.0	1730	22	AAE20499
64	25	96.2	99	21	AAV49808
65	25	96.2	111	23	AAV00690
66	25	96.2	139	18	AAW23003
67	25	96.2	139	19	AAW72663
68	25	96.2	139	22	AAE51320
69	25	96.2	150	21	AAV49807
70	25	96.2	156	21	AAV11977
71	25	96.2	168	21	AAV49806
72	25	96.2	174	21	AAV11976
73	25	96.2	219	21	AAV42151
74	25	96.2	259	21	AAV49826
75	25	96.2	287	21	AAV36626

ALIGNMENTS

RESULT 1
AAV17156
ID AAV17156 standard; peptide; 5 AA.

XX AAY17156;
 XX 20-SEP-1999 (first entry)
 XX Cadherin-mediated adhesion modulating cyclic peptide.
 XX
 XX Cell adhesion modulation; CAM; synaptic stability; cadherin;
 KW cadherin-mediated adhesion; drug delivery; cell adhesion; tumour;
 KW wound healing; neurite outgrowth; cyclic; circular.
 XX
 XX Synthetic.
 XX
 XX WO9933875-A1.
 XX 08-JUL-1999.
 XX
 XX 23-DEC-1998; 98WO-CA01207.
 XX
 XX 23-DEC-1997; 97US-0996679.
 XX (UTMC-) UNIV MCGILL.
 XX
 XX Blaschuk OW, Gour BJ;
 XX
 XX WPI; 1999-430231/36.
 XX
 XX Cyclic peptide cell adhesion modulating agents, useful for
 PT modulating synaptic stability
 XX
 XX Claim 17; Page 65; 144pp; English.
 XX
 XX The invention provides cyclic peptide cell adhesion modulating (CAM)
 CC agents that comprises a His-Ala-Val recognition sequence. Also provided
 CC is a method for inhibiting synaptic stability in a mammal that comprises
 CC administering to a mammal a therapeutically effective amount of a CAM
 CC agent that inhibits cadherin-mediated adhesion, where the agent comprises
 CC a cyclic peptide having a peptide ring, and where the sequence His-Ala-
 CC Val is present within the peptide ring. The cyclic peptides are cell
 CC adhesion modulating agents that inhibit cadherin-mediated adhesion. They
 CC can be used in a method for inhibiting synaptic stability in mammals. The
 CC agents can be used to treat diseases or other conditions characterized by
 CC undesirable cell adhesion or to facilitate drug delivery to a specific
 CC tissue or tumour. Alternatively the agents may be used to enhance cell
 CC adhesion (e.g. to supplement or replace stitches or to facilitate wound
 CC healing) or to enhance or direct neurite outgrowth.
 XX
 XX SQ Sequence 5 AA;
 Query Match 100.0%; Score 26; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVE 5
 Db |||||
 1 KHAVE 5
 RESULT 2
 AAW91035
 ID AAW91035 standard; peptide; 5 AA.
 XX
 XX AC AAW91035;
 XX
 XX 24-MAR-1999 (first entry)
 XX
 XX Peptide inhibiting interaction of alpha-catenin and beta-catenin.
 XX
 XX Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;
 KW cadherin-mediated function; demyelinating neurological disease;
 KW multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;
 KW central nervous system; apoptosis induction; cadherin-expression cell;
 KW pregnancy prevention; vasopermeability; synaptic stability; diabetes;

KW rheumatoid arthritis; allergic response; learning; memory;
 KW antibody-mediated graft rejection; cyclic.
 XX
 XX Synthetic.
 XX
 XX WO9845319-A2.
 XX
 XX 15-OCT-1998.
 XX
 XX 14-APR-1998; 98WO-CA00322.
 XX
 XX 10-APR-1997; 97US-0043361.
 XX
 XX (UTMC-) UNIV MCGILL.
 XX
 XX Blaschuk OW, Gour BJ;
 XX
 XX WPI; 1999-024009/02.
 XX
 XX New catenin modulating agents - comprising peptides having a
 PT sequence HAV or analogues or antibodies, used for modulating
 PT cadherin-mediated functions
 XX
 XX Claim 12; Page 78; 106pp; English.
 XX
 XX The present sequence represents a peptide which is capable of inhibiting
 CC an interaction between alpha-catenin and beta-catenin. The peptide is
 CC used in modulating agents that are used for modulating cadherin-mediated
 CC functions. They can be used for disrupting interaction between
 CC alpha-catenin and beta-catenin in a cell inhibiting cell adhesion, e.g.
 CC between epithelial cells, endothelial cells, neural cells, tumour cells
 CC and lymphocytes, for treating a demyelinating neurological disease, e.g.
 CC multiple sclerosis, for reducing unwanted cellular adhesion in a mammal,
 CC for enhancing the delivery of a drug through the skin of a mammal, for
 CC enhancing the delivery of a drug to a tumour in a mammal, for treating
 CC cancer in a mammal, for inhibiting angiogenesis in a mammal, for
 CC enhancing drug delivery to the central nervous system of a mammal, for
 CC inducing apoptosis in a cadherin-expression cell, for modulating the
 CC immune system of a mammal, for preventing pregnancy in a mammal, for
 CC increasing vasopermeability in a mammal, or for inhibiting synaptic
 CC stability in a mammal. In particular they can be used for treating
 CC diabetes, rheumatoid arthritis, allergic responses, antibody-mediated
 CC graft rejection or for stimulating learning and memory.
 XX
 XX SQ Sequence 5 AA;
 Query Match 100.0%; Score 26; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVE 5
 Db |||||
 1 KHAVE 5
 RESULT 3
 AAB27348
 ID AAB27348 standard; peptide; 5 AA.
 XX
 XX AC AAB27348;
 XX
 XX 15-FEB-2001 (first entry)
 XX
 XX Beta-catenin derived cyclic peptide SEQ ID NO: 36.
 XX
 XX Beta-catenin; cadherin-mediated intercellular adhesion;
 KW cell differentiation; modulating agent; hair loss; skin exfoliation;
 KW internalisation moiety; flanking sequence; transcription; hearing loss;
 KW cyclic.
 XX
 XX Unidentified.
 XX
 XX Key Location/Qualifiers

FT Misc-difference 1...5
 XX /note= "the molecule is cyclised"

PN WO200053632-A1.

XX 14-SEP-2000.

XX 07-MAR-2000; 2000WO-CA00222.

XX 09-MAR-1999; 99US-0265107.

XX (UTMC-) UNIV MCGILL.

XX Blaschuk OW, Gour BJ;

XX WPI; 2000-594308/56.

XX Stimulating beta-catenin mediated gene expression, cellular
 XX differentiation and hair growth, involves contacting cells with
 XX modulating agent capable of inhibiting interaction between alpha and
 XX beta catenin

XX Disclosure; Page 7; 77pp; English.

CC The present invention is concerned with methods of modulating the amount
 CC of free beta-catenin in the cell, and methods of stimulating the
 CC expression of genes involved in cellular differentiation, the
 CC transcription of which is under the control of beta-catenin. The
 CC peptides given in AAB27083-B27088, AAB27284-B27300 and AAB27330-B27351
 CC can be used as modulating agents which interrupt the interaction between
 CC alpha and beta catenin, causing increased levels of the latter and
 CC stimulating the activation of beta-catenin mediated transcription. This
 CC can be used to stimulate cell differentiation, which can then be used to
 CC promote hair growth and skin exfoliation. This latter is particularly
 CC useful in the improvement of photodamaged skin and to minimise wrinkles.
 CC The modulating peptide can also be used to reduce hearing loss resulting
 CC from inner ear disorders such as hyperacusis and tinnitus.

XX Sequence 5 AA;

Query Match 100.0%; Score 26; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5

DB 1 KHAVE 5

RESULT 4

AAV73805

ID AAV73805 standard; Peptide; 5 AA.

XX AAV73805;

XX 07-MAR-2000 (first entry)

XX N-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:313.

XX Diagnosis; cancer; OB-cadherin; N-cadherin; E-cadherin; CAR; cyclic;
 XX cell adhesion recognition sequence; leukaemia; metastasis;
 XX prostate cancer; breast cancer; ovarian cancer; carcinoma.

XX Synthetic.

OS Homo sapiens.

XX WO9957565-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00362.

XX 05-MAY-1998; 98US-0073040.

PR 06-NOV-1998; 98US-0187859.
 PR 20-JAN-1999; 99US-0234395.

PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-062165/05.

XX Detecting expression of OB-cadherin or N-cadherin used for diagnosing
 XX and evaluating cancer

XX Disclosure; Page 22; 56pp; English.

CC The present invention describes methods which have been developed for
 CC detecting and evaluating cancer that are based on the finding that
 CC OB-cadherin and N-cadherin are expressed by metastatic carcinoma cells,
 CC but not by highly differentiated, poorly invasive carcinomas. A method
 CC for determining the presence or absence of a cancer in a patient
 CC comprises: (1) contacting a biological sample from the patient with a
 CC binding agent that specifically binds to OB- or N-cadherin OR with an
 CC oligonucleotide that hybridises to a polynucleotide that encodes OB- or
 CC N-cadherin; and (2) detecting in the sample an amount of polypeptide
 CC that binds to the binding agent OR the amount of polynucleotide that
 CC hybridises to the oligonucleotide, relative to a predetermined cut-off
 CC value, and determining the presence or absence of cancer in the patient
 CC from this. The methods from the present invention can be used to
 CC determine the metastatic potential of a cancer. The methods may be used
 CC to detect a metastatic cancer in a patient, to monitor progression of a
 CC cancer, or to evaluate the metastatic potential of a cancer. Cancers
 CC which may be evaluated using the methods include leukemia, prostate
 CC cancer, breast cancer and ovarian cancer. AAV73501 to AAV73813 represent
 CC peptide sequences used in the exemplification of the present invention.
 CC Specifically, AAV73503 to AAV73789 represent OB-cadherin cell adhesion
 CC recognition (CAR) peptides and AAV73790 to AAV73808 represent N-cadherin
 CC CAR peptides. AA247906 to AA247915 represent PCR primers used in examples
 CC from the present invention.

XX Sequence 5 AA;

Query Match 100.0%; Score 26; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5

DB 1 KHAVE 5

RESULT 5

AA65376

ID AAG65376 standard; peptide; 5 AA.

XX AAG65376;

XX 30-NOV-2001 (first entry)

XX Cyclic peptide with cadherin CAR sequence.

XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 XX cytostatic; vulnery; immunomodulator; vasotropic; neuroprotective;
 XX cerebroprotective; muscular; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

XX Modified-site 1

XX /note= "putative N-terminal acetylation or
 XX alkoxybenzylation"

XX Modified-site 5

XX /note= "putative C-terminal amide or ester"

XX WO200153331-A2.

XX PD 26-JUL-2001.
XX PF 24-JAN-2001; 2001WO-US02508.
XX PR 24-JAN-2000; 2000US-0491078.
XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX PI Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
XX PT Hu Z;
XX DR WPI; 2001-549899/61.
XX PT Cell adhesion modulating agent used for enhancing delivery of drug to
XX PT tumor comprises imidazole compounds -
XX PS Examples; Fig 3A-I; 436pp; English.
XX CC The invention relates to cell adhesion modulating agents that comprise
XX CC imidazole compounds of specified formulae that are peptidomimetics of
XX CC cyclic peptides. The peptidomimetics have a structure similar to that of
XX CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
XX CC The agents are used for modulating classical cadherin mediated
XX CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
XX CC delivery of a drug to a tumor, inhibiting development of cancer,
XX CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
XX CC wound healing, modulating the immune system, increasing vasopermeability,
XX CC treating demyelinating disease, facilitating migration of an N-cadherin
XX CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
XX CC neurite outgrowth, and treating spinal cord injuries and macular
XX CC degeneration. The present sequence represents a cyclic peptide with
XX CC cadherin cell adhesion recognition (CAR) sequence.
XX SQ Sequence 5 AA;
Query Match 100.0%; Score 26; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05; Mismatches 0; Gaps 0;
Matches 5; Conservative 0; Indels 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
DB 1 KHAVE 5
RESULT 6
AAM47483
ID AAM47483 standard; peptide; 5 AA.
XX AC AAM47483;
XX DT 12-FEB-2002 (first entry)
XX DE Cyclic peptide endothelial cell adhesion modulator #4.
XX KW Cadherin; cytostatic; gynecological; endometriosis;
XX KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
XX KW cyclic.
XX OS Synthetic.
XX PH Key Location/Qualifiers
FT Modified-site 1 /note= "Linked to residue 5 to form a cyclic peptide,
FT N-terminal acetyl or N-terminal CH3-SO2 group"
FT Modified-site 5 /note= "Linked to residue 1 to form a cyclic peptide,
FT C-terminal amide"
XX PN WO200177146-A2.
XX PD 18-OCT-2001.
XX PT

PF 09-APR-2001; 2001WO-US11669.
XX PR 07-APR-2000; 2000US-0544782.
XX PA (UYMC-) UNIV MCGILL.
XX PI Blaschuk OW, Gour BJ, Farookhi R, Ali A;
XX DR WPI; 2002-049129/06.
XX CC Modulating endothelial cell adhesion for inhibiting development of
XX CC endometriosis, increasing blood flow to tumor in a mammal, by
XX CC contacting cell with a cyclic peptide having cadherin cell adhesion
XX CC recognition sequence
XX PS Claim 6; Page 64; 139pp; English.
XX CC The present invention relates to a method for modulating endothelial cell
XX CC adhesion. The method comprises contacting an endothelial cell with a
XX CC modulating peptide comprising a cadherin cell adhesion recognition
XX CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
XX CC for inhibiting angiogenesis, increasing vasopermeability, increasing
XX CC blood flow to a tumour, disrupting neovasculture and inhibiting the
XX CC development of endometriosis in a mammal. The modulating peptide reduces
XX CC unwanted endothelial adhesion occurring between tumour cells, tumour
XX CC cells and normal cells, normal cells as a result of surgery, injury,
XX CC chemotherapy, disease and inflammation. The present sequence is one such
XX CC modulating peptide.
XX SQ Sequence 5 AA;
Query Match 100.0%; Score 26; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05; Mismatches 0; Gaps 0;
Matches 5; Conservative 0; Indels 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
DB 1 KHAVE 5
RESULT 7
ABG01734
ID ABG01734 standard; Protein; 31 AA.
XX AC ABG01734;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #1725.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0849167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS65921.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID No 32093; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 31 AA;

Query Match 100.0%; Score 26; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
 |||||
 Db 3 KHAVE 7

RESULT 8
 AAB01735
 ID ABG01735 standard; Protein; 213 AA.
 XX
 AC AAB01735;
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #1726.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang Yt;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS65922.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX
 PS Claim 20; SEQ ID No 32094; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 213 AA;

Query Match 100.0%; Score 26; DB 22; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
 |||||
 Db 129 KHAVE 133

RESULT 9
 AAB47101
 ID AAB47101 standard; Protein; 266 AA.
 XX
 AC AAB47101;
 DT 16-MAY-2001 (first entry)
 XX
 DE B. methanolicus dihydrodipicolinate reductase.
 XX
 KW Thermophilic; Bacillus methanolicus; dihydrodipicolinate synthase;
 KW DDPS; dihydrodipicolinate reductase; DDPR; L-lysine.
 XX
 OS Bacillus methanolicus.
 XX
 PN EP1074626-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 01-AUG-2000; 2000EP-0116623.
 XX
 PR 04-AUG-1999; 99JP-0221468.
 XX
 PA (AJIN) AJINOMOTO KK.
 XX
 PI Tsujimoto N, Yasueda H, Kawahara Y, Sugimoto S;
 XX
 DR WPI; 2001-171048/18.
 DR N-PSDB; AAC85499.
 XX
 PT Bacillus methanolicus dihydrodipicolinate synthase and
 PT dihydrodipicolinate reductase proteins, useful for breeding L-lysine
 PT producing bacteria -
 XX
 PS Claim 5; Page 15-16; 18pp; English.
 XX
 CC The sequences given in AAB47100 and AAB47101 represent the

CC thermophilic Bacillus methanolicus dihydrodipicolinate synthase (DDPS)
 CC and dihydrodipicolinate reductase (DDPR) proteins, respectively. The
 CC DDPS and DDPR genes are useful for producing DDPS and DDPR, and
 CC for breeding L-lysine producing bacteria. These DDPS and DDPR genes
 CC allow L-lysine to be produced at elevated temperatures.

SQ Sequence 266 AA;
 Query Match 100.0%; Score 26; DB 22; Length 266;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
 DB 245 KHAVE 249
 |||||

RESULT 10
 AAB99734
 ID AAB99734 standard; Protein, 332 AA.
 XX
 AC AAB99734;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Oryza sativa peroxidase r1420 protein SEQ ID NO:6.
 XX
 KW Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;
 KW modification; plant; bacterial infection; Magnaporthe grisea.
 XX
 OS Oryza sativa.
 XX
 FN WO200142475-A1.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000WO-JP08728.
 XX
 PR 10-DEC-1999; 99JP-0352472.
 XX
 PA (NORQ) JAPAN MIN AGRIC FORESTRY & FISHERIES.
 XX
 PI Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;
 PI Hiraga S;
 DR WPI; 2001-381695/40.
 DR N-PSDB; AAB44073.
 XX
 PT New set of rice peroxidase genes for analysis of peroxidase expression
 PT in rice under varying conditions and production of rice plants with
 PT desired characteristics -
 XX
 PS Disclosure; Page 131-133; 258pp; Japanese.

CC The present invention describes a set of peroxidase genes found in
 CC plants, especially rice, and their homologues, modified forms and
 CC fragments, where the sequences of the peroxidase genes in the set are
 CC given in AAB44071 to AAB44091. Also described are: (1) promoters for the
 CC control of the gene set; (2) the preparation of cassette vectors using
 CC the genes and promoters; (3) analysis of plant characteristics using the
 CC peroxidase set by isolating RNA from the plant, binding the RNA to a
 CC membrane, mixing with a labelled peroxidase gene set, incubating, and
 CC detecting the label signal to show which genes in the set are expressed
 CC in the sample plant; and (4) DNA microarrays for peroxidase gene
 CC expression analysis. The set of genes are used for the analysis of the
 CC pattern of peroxidase gene expression in particular rice plants and
 CC their component tissues and under different environmental conditions,
 CC and modification of rice plants to provide desired specificities of
 CC peroxidase gene expression to impart particular characteristics to the
 CC plants such as response to bacterial infection by Magnaporthe grisea.
 CC The present sequence represents a rice peroxidase encoded by a gene from
 CC the gene set described above.

SQ Sequence 332 AA;
 Query Match 100.0%; Score 26; DB 22; Length 332;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
 DB 113 KHAVE 117
 |||||

RESULT 11
 AAY54091
 ID AAY54091 standard; Protein, 366 AA.
 XX
 AC AAY54091;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Enzyme EPSF involved in exopolysaccharide biosynthesis.
 XX
 KW Exopolysaccharide; EPS; ESP enzyme; EPSA; EPSB; EPSD; EPSH; EPSG; EPSF;
 KW EPSH; EPSJ; EPSU; EPSK; Lactobacillus delbrueckii bulgaricus;
 KW activated D-galactose pyranose; saccharide; beta-glycosyltransferase;
 KW transcription attenuator; glucosyl-phosphotransferase;
 KW alpha-glycosyltransferase; glucosyltransferase; EPSL; EPSM;
 KW alpha-glycosyltransferase; EPS polymerase; glycosyltransferase; EPSN;
 KW phosphotransferase; transporter; food; fermented milk product; yoghurt;
 KW cheese; flavour stability; organoleptic property.
 XX
 OS Lactobacillus delbrueckii bulgaricus.
 XX
 PN WO9962316-A2.
 XX
 PD 09-DEC-1999.
 XX
 PF 22-APR-1999; 99WO-EP02841.
 XX
 PR 22-APR-1998; 98EP-0201310.
 PR 22-APR-1998; 98EP-0201311.
 PR 22-APR-1998; 98EP-0201312.
 XX
 PA (NEST) SOC. PROD. NESTLE SA.
 XX
 PI Stangele F, Germond JB, Lamothe G;
 XX
 DR WPI; 2000-097267/08.
 DR N-PSDB; AAZ45260.
 XX
 PT New recombinant enzymes for synthesis of exopolysaccharides,
 PT particularly in lactic acid bacteria, for improving properties of
 PT fermented milk products -
 XX
 PS Claim 13; Page 152; 162pp; French.

CC AAY54086-99 represent enzymes involved in the biosynthesis of
 CC exopolysaccharides (EPS). These enzymes are designated EPSA-EPSN, and
 CC are encoded by open reading frames epsa-epsn. The enzymes are isolated
 CC from Lactobacillus delbrueckii bulgaricus. The proteins are used
 CC in a method for the synthesis of EPS, which includes at least one step
 CC of forming a bond (alpha or beta-isomer) between C-1 (carrying the
 CC reducing aldehyde function, of an activated D-galactose pyranose), and
 CC a phosphate on a lipophilic or proteinaceous primer. Stepwise synthesis
 CC of EPS occurs with, in each step, addition of a new sugar unit, through
 CC its hemi-acetyl function, to an alcoholic hydroxyl of a second sugar
 CC unit, present at the end of a chain of sugar residues bonded to the
 CC primer. EPSA is attenuator of transcription which influences
 CC regulation; EPSB determines the length of the saccharide chain; EPSG is
 CC involved in the control of the molecular weight and/or the length of the
 CC polysaccharide; EPSD, EPSL and EPSM are involved in synthesis of EPS;
 CC EPSH is a galactosyl- or glucosyl-phosphotransferase which catalyses the
 CC transfer of the first saccharide on the primer; EPSF and EPSG are
 CC alpha-glycosyltransferases; EPSH and EPSHJ are

CC beta-glucosyltransferases; EPSG is a glucosyltransferase; EPSK is
 CC responsible for the polymerisation of the repetitive units; and EPSN is
 CC responsible for export of the EPS. The EPS enzyme are used to improve
 CC properties of foods, particularly fermented milk products such as yoghurt
 CC and cheese, e.g. their organoleptic properties and flavour stability.

XX Sequence 366 AA;
 SQ Query Match 100.0%; Score 26; DB 21; Length 366;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAIVE 5
 Db 169 KHAIVE 173

RESULT 12
 AAY43793
 ID AAY43793 standard; Protein; 366 AA.

XX AAY43793;
 AC AAY43793;
 XX 11-FEB-2000 (first entry)
 DT 11-FEB-2000 (first entry)
 DE Amino acid sequence of epsF of L. delbrueckii bulgaricus Lfi5.
 XX
 KW eps operon; Lactobacillus delbrueckii bulgaricus Lfi5; enzyme; epsA;
 KW epsB; epsC; epsD; epsE; epsF; epsG; epsH; epsI; epsJ; epsK; epsL; epsM;
 KW exopolysaccharide biosynthesis; EPS; intersugar bond; antitumor;
 KW probiotic; foodstuff; organoleptic quality; flavour;
 KW lactic acid bacteria; acidified milk product; yoghurt; cheese.

XX Lactobacillus delbrueckii bulgaricus.

OS W09554475-A2.

XX 28-OCT-1999.

XX 22-APR-1999; 99WO-EP03011.

XX 22-APR-1998; 98EP-0201310.

XX 22-APR-1998; 98EP-0201311.

XX 22-APR-1998; 98EP-0201312.

XX (NEST) SOC PROD NESTLE SA.

XX Stingele F, Germond JE, Lamothe G;

XX WPI: 2000-013255/01.

XX N-PSDB; AA230357, AAY43796, AAY43797, AAY43798, AAY43799, AAY43800,

XX AAY43801, AAY43842.

XX New recombinant enzymes for biosynthesis of exopolysaccharides having
 PT e.g. antitumor or probiotic properties or useful in fermented milk
 PT products -

XX Claim 13; Page 153; 163pp; French.

XX AAY43788-89, AAY43791-Y437801 and AAY43842 represent the enzymes encoded
 CC by the eps operon of Lactobacillus delbrueckii bulgaricus Lfi5. The
 CC operon contains 14 open reading frames, and encodes enzymes (epsA,
 CC epsB, epsC, epsD, epsE, epsF, epsG, epsH, epsI and epsJ, epsK, epsL,
 CC and epsM) that are involved in the biosynthesis of exopolysaccharides
 CC (EPS). The enzymes catalyse the formation of specific intersugar bonds.
 CC The enzymes catalyse a process that includes at least one step of
 CC forming a bond (in alpha or beta anomeric form) between C1, carrying
 CC the reducing aldehyde group of an activated D-Galp (galactose in
 CC pyranose form), and a phosphate group on a lipophilic or proteinaceous
 CC primer. The enzymes are used to produce EPS that have antitumor or
 CC probiotic properties or are used in foodstuffs to improve organoleptic
 CC qualities and flavour. When expressed by lactic acid bacteria, EPS
 CC impart a free-flowing character and/or a smooth, creamy texture to

CC acidified milk products (yoghurt or cheese).

XX Sequence 366 AA;

SQ Query Match 100.0%; Score 26; DB 21; Length 366;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAIVE 5
 Db 169 KHAIVE 173

RESULT 13
 AAE21827
 ID AAE21827 standard; Protein; 377 AA.

XX AAE21827;

XX 16-JUL-2002 (first entry)

DE Human 46873 asparaginase domain consensus.

XX Human; asparaginase; haematopoietic neoplastic disorder; hyperplastic;
 KW neoplastic; lymphoid malignancy; acute lymphoblastic leukaemia; tumour;
 KW Hodgkin's disease; Reed-Sternberg disease; lymphosarcoma; gene therapy.

XX Homo sapiens.

OS US2002038014-A1.

XX 28-MAR-2002.

XX 23-MAR-2001; 200IUS-0816664.

XX 24-MAR-2000; 2000US-191973P.

XX (MEYE/) MEYERS R A.

XX Meyers RA;

XX WPI: 2002-329129/36.

XX Asparaginase 26443 and 4687 enzymes and nucleic acids, useful in gene
 PT therapy and diagnosis -

XX Disclosure; Fig 8; 58pp; English.

XX The invention relates to nucleic acid molecules, designated 26443 and
 CC 46873, which encode novel asparaginases. The polynucleotides and
 CC polypeptides of the invention may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate asparaginase
 CC expression such as haematopoietic neoplastic disorders which include
 CC diseases involving hyperplastic/neoplastic cells of haematopoietic
 CC origin, lymphoid malignancies such as acute lymphoblastic leukaemia,
 CC Hodgkin's disease and Reed-Sternberg disease, lymphosarcoma and other
 CC forms of solid tumours. The polynucleotides of the invention are used
 CC in gene therapy and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acids in samples, and therefore which patients may be
 CC in need of restorative therapy. The asparaginases may also be used as
 CC antigens in the production of antibodies against it and the antibodies,
 CC agonists and antagonists may also be used to regulate its expression and
 CC activity. The antibodies may also be used as diagnostic agents for
 CC detecting the presence of asparaginases in samples (e.g. by enzyme linked
 CC immunosorbent assay (ELISA)). The present sequence is human 46873
 CC asparaginase domain consensus protein.

XX Sequence 377 AA;

SQ Query Match 100.0%; Score 26; DB 23; Length 377;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 38 KHAVE 42

RESULT 14

AAE21814
ID AAE21814 standard; Protein; 378 AA.

XX AAE21814;

XX 16-JUL-2002 (first entry)

XX Human 26443 asparaginase domain consensus.

XX Human; asparaginase; haematopoietic neoplastic disorder; hyperplastic;
XX neoplastic; lymphoid malignancy; acute lymphoblastic leukaemia; tumour;
XX Hodgkin's disease; Reed-Sternberg disease; lymphosarcoma; gene therapy.

XX Homo sapiens.

XX US2002038014-A1.

XX 28-MAR-2002.

XX 23-MAR-2001; 2001US-0816664.

XX 24-MAR-2000; 2000US-191973P.

XX (MEYE/) MEYERS R A.

XX Meyers RA;

XX WPI; 2002-329129/36.

XX Asparaginase 26443 and 4687 enzymes and nucleic acids, useful in gene
XX therapy and diagnosis -

XX Disclosure; Fig 4; 58pp; English.

XX The invention relates to nucleic acid molecules, designated 26443 and
XX 46873, which encode novel asparaginases. The polynucleotides and
XX polypeptides of the invention may be used in the prevention, diagnosis
XX and treatment of diseases associated with inappropriate asparaginase
XX expression such as haematopoietic neoplastic disorders which include
XX diseases involving hyperplastic/neoplastic cells of haematopoietic
XX origin, lymphoid malignancies such as acute lymphoblastic leukaemia,
XX Hodgkin's disease and Reed-Sternberg disease, lymphosarcoma and other
XX forms of solid tumours. The polynucleotides of the invention are used
XX in gene therapy and its complementary sequences may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acids in samples, and therefore which patients may be
XX in need of restorative therapy. The asparaginases may also be used as
XX antigens in the production of antibodies against it and the antibodies,
XX agonists and antagonists may also be used to regulate its expression and
XX activity. The antibodies may also be used as diagnostic agents for
XX detecting the presence of asparaginases in samples (e.g. by enzyme linked
XX immunosorbent assay (ELISA)). The present sequence is human 26443
XX asparaginase domain consensus protein.

XX Sequence 378 AA;

Query Match 100.0%; Score 26; DB 23; Length 378;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 39 KHAVE 43

RESULT 15

AAE21814

AAE21814

ID AAE21814 standard; Protein; 378 AA.

AAE21814

ID AAE21814 standard; Protein; 378 AA.

XX AAE21814;

XX 16-JUL-2002 (first entry)

XX Human 26443 asparaginase domain consensus.

XX Human; asparaginase; haematopoietic neoplastic disorder; hyperplastic;
XX neoplastic; lymphoid malignancy; acute lymphoblastic leukaemia; tumour;
XX Hodgkin's disease; Reed-Sternberg disease; lymphosarcoma; gene therapy.

XX Homo sapiens.

XX US2002038014-A1.

XX 28-MAR-2002.

XX 23-MAR-2001; 2001US-0816664.

XX 24-MAR-2000; 2000US-191973P.

XX (MEYE/) MEYERS R A.

XX Meyers RA;

XX WPI; 2002-329129/36.

XX Asparaginase 26443 and 4687 enzymes and nucleic acids, useful in gene
XX therapy and diagnosis -

XX Disclosure; Fig 4; 58pp; English.

XX The invention relates to nucleic acid molecules, designated 26443 and
XX 46873, which encode novel asparaginases. The polynucleotides and
XX polypeptides of the invention may be used in the prevention, diagnosis
XX and treatment of diseases associated with inappropriate asparaginase
XX expression such as haematopoietic neoplastic disorders which include
XX diseases involving hyperplastic/neoplastic cells of haematopoietic
XX origin, lymphoid malignancies such as acute lymphoblastic leukaemia,
XX Hodgkin's disease and Reed-Sternberg disease, lymphosarcoma and other
XX forms of solid tumours. The polynucleotides of the invention are used
XX in gene therapy and its complementary sequences may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acids in samples, and therefore which patients may be
XX in need of restorative therapy. The asparaginases may also be used as
XX antigens in the production of antibodies against it and the antibodies,
XX agonists and antagonists may also be used to regulate its expression and
XX activity. The antibodies may also be used as diagnostic agents for
XX detecting the presence of asparaginases in samples (e.g. by enzyme linked
XX immunosorbent assay (ELISA)). The present sequence is human 26443
XX asparaginase domain consensus protein.

XX Sequence 378 AA;

Query Match 100.0%; Score 26; DB 23; Length 378;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 39 KHAVE 43

AAE21814

ID AAE21814 standard; Protein; 378 AA.

XX AAE21814;

XX 16-JUL-2002 (first entry)

XX Human 26443 asparaginase domain consensus.

XX Human; asparaginase; haematopoietic neoplastic disorder; hyperplastic;
XX neoplastic; lymphoid malignancy; acute lymphoblastic leukaemia; tumour;
XX Hodgkin's disease; Reed-Sternberg disease; lymphosarcoma; gene therapy.

XX Homo sapiens.

XX US2002038014-A1.

XX 28-MAR-2002.

XX 23-MAR-2001; 2001US-0816664.

XX 24-MAR-2000; 2000US-191973P.

XX (MEYE/) MEYERS R A.

XX Meyers RA;

XX WPI; 2002-329129/36.

XX Asparaginase 26443 and 4687 enzymes and nucleic acids, useful in gene
XX therapy and diagnosis -

XX Disclosure; Fig 4; 58pp; English.

XX The invention relates to nucleic acid molecules, designated 26443 and
XX 46873, which encode novel asparaginases. The polynucleotides and
XX polypeptides of the invention may be used in the prevention, diagnosis
XX and treatment of diseases associated with inappropriate asparaginase
XX expression such as haematopoietic neoplastic disorders which include
XX diseases involving hyperplastic/neoplastic cells of haematopoietic
XX origin, lymphoid malignancies such as acute lymphoblastic leukaemia,
XX Hodgkin's disease and Reed-Sternberg disease, lymphosarcoma and other
XX forms of solid tumours. The polynucleotides of the invention are used
XX in gene therapy and its complementary sequences may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acids in samples, and therefore which patients may be
XX in need of restorative therapy. The asparaginases may also be used as
XX antigens in the production of antibodies against it and the antibodies,
XX agonists and antagonists may also be used to regulate its expression and
XX activity. The antibodies may also be used as diagnostic agents for
XX detecting the presence of asparaginases in samples (e.g. by enzyme linked
XX immunosorbent assay (ELISA)). The present sequence is human 26443
XX asparaginase domain consensus protein.

XX Sequence 378 AA;

Query Match 100.0%; Score 26; DB 23; Length 378;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 39 KHAVE 43

XX AC AAY74538;
 XX DT 21-MAR-2000 (first entry)
 XX DE
 XX DE Neisseria meningitidis ORF 140 protein sequence SEQ ID NO:550.
 XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX OS Neisseria meningitidis.
 XX PN WO9957280-A2.
 XX PD 11-NOV-1999.
 XX PF 30-APR-1999; 99WO-US09346.
 XX PR 01-MAY-1998; 98US-0083758.
 XX PR 31-JUL-1998; 98US-0094869.
 XX PR 02-SEP-1998; 98US-0098994.
 XX PR 02-SEP-1998; 98US-0099062.
 XX PR 09-OCT-1998; 98US-0103749.
 XX PR 09-OCT-1998; 98US-0103794.
 XX PR 09-OCT-1998; 98US-0103796.
 XX PR 25-FEB-1999; 99US-0121528.
 XX PA (CHIR) CHIRON CORP.
 XX PA (GENO-) INST GENOMIC RES.
 XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX DR WPI; 2000-062150/05.
 XX DR N-PSDB; AA253300.
 XX PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX PS Claim 2; Page 395; 1453pp; English.
 XX CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisseria bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX SQ Sequence 454 AA;
 Query Match 100.0%; Score 26; DB 21; Length 454;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVE 5
 Db 85 KHAVE 89
 RESULT 17
 AAY74539
 ID AAY74539 standard; Protein; 454 AA.
 XX AC AAY74539;

XX DT 21-MAR-2000 (first entry)
 XX DE
 XX DE Neisseria meningitidis ORF 140 protein sequence SEQ ID NO:552.
 XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX OS Neisseria meningitidis.
 XX PN WO9957280-A2.
 XX PD 11-NOV-1999.
 XX PF 30-APR-1999; 99WO-US09346.
 XX PR 01-MAY-1998; 98US-0083758.
 XX PR 31-JUL-1998; 98US-0094869.
 XX PR 02-SEP-1998; 98US-0098994.
 XX PR 02-SEP-1998; 98US-0099062.
 XX PR 09-OCT-1998; 98US-0103749.
 XX PR 09-OCT-1998; 98US-0103794.
 XX PR 09-OCT-1998; 98US-0103796.
 XX PR 25-FEB-1999; 99US-0121528.
 XX PA (CHIR) CHIRON CORP.
 XX PA (GENO-) INST GENOMIC RES.
 XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX DR WPI; 2000-062150/05.
 XX DR N-PSDB; AA253301.
 XX PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX PS Claim 2; Page 397; 1453pp; English.
 XX CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisseria bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX SQ Sequence 454 AA;
 Query Match 100.0%; Score 26; DB 21; Length 454;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVE 5
 Db 85 KHAVE 89
 RESULT 18
 AAB25591
 ID AAB25591 standard; Protein; 461 AA.
 XX AC AAB25591;
 XX DT 21-NOV-2000 (first entry)

XX DE Protein encoded by human secreted protein gene #9 clone HTTDB46.

XX KW Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;

XX KW antirheumatic, dermatological; antiproliferative; antiarteriosclerotic;

XX KW anticancer; vulnary; antiviral; antibacterial; antifungal;

XX KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;

XX KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;

XX KW Crohn's disease; nephritis; hyperproliferative disorder;

XX KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;

XX KW melanoma; lymphoma; wound healing; human.

XX OS Homo sapiens.

XX PN WO200029435-A1.

XX PD 25-MAY-2000.

XX PF 27-OCT-1999; 99WO-US25031.

XX PR 28-OCT-1998; 98US-0105971.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;

XX PI Greene JM;

XX DR WPI; 2000-387742/33.

XX KW Isolated nucleic acid molecules encoding human secreted proteins are

XX KW used for the prevention, amelioration and treatment of autoimmune,

XX KW inflammatory, hyperproliferative and cardiovascular disorders, cancer,

XX KW PT wounds, and infectious diseases -

XX PS Claim 1; Page 682-683; 803pp; English.

XX CC The present invention relates to 12 secreted human proteins and the

XX CC nucleotide sequences encoding them. The polynucleotide sequences given

XX CC in AAA80606-A80623 encode the 12 secreted protein sequences given in

XX CC AAA25576-B25593. The human secreted proteins have various activities

XX CC dependent on the tissues in which they are expressed. Examples of the

XX CC activities of the proteins include: immunosuppressant;

XX CC anti-inflammatory; antiarthritic; antirheumatic, dermatological;

XX CC antiviral; antibacterial; and antifungal activity. The proteins,

XX CC and/or agonists and antagonists may be used to treat prevent

XX CC and/or diagnose various disease, disorders and conditions examples of

XX CC arthritis, dermatitis, and multiple sclerosis; inflammatory disorders

XX CC e.g. inflammatory bowel disease, Crohn's disease and nephritis;

XX CC hyperproliferative disorders such as paraproteinemia and purpura;

XX CC cardiovascular disorders e.g. coronary arteriosclerosis and myocarditis;

XX CC cancer e.g. melanoma and lymphoma. The proteins and polynucleotide

XX CC sequences may also be used in wound healing and the treatment of

XX CC infectious diseases. The human secreted protein gene #9 and protein

XX CC sequences are represented in sequences AAA80614 and AAA25584. Sequences

XX CC AAA80664-A80668 represent genes related to the secreted protein gene#9.

XX SQ Sequence 461 AA;

Query Match 100.0%; Score 26; DB 21; Length 461;

Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5

Db 285 KHAVE 289

|||||

RESULT 19

AA255615

ID AAB25615 standard; Protein; 461 AA.

XX

AC AAB25615;

XX 21-NOV-2000 (first entry)

DT Human secreted protein #9.

XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;

XX KW antirheumatic, dermatological; antiproliferative; antiarteriosclerotic;

XX KW anticancer; vulnary; antiviral; antibacterial; antifungal;

XX KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;

XX KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;

XX KW Crohn's disease; nephritis; hyperproliferative disorder;

XX KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;

XX KW melanoma; lymphoma; wound healing; human.

XX OS Homo sapiens.

XX PN WO200029435-A1.

XX PD 25-MAY-2000.

XX PF 27-OCT-1999; 99WO-US25031.

XX PR 28-OCT-1998; 98US-0105971.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;

XX PI Greene JM;

XX DR WPI; 2000-387742/33.

XX KW Isolated nucleic acid molecules encoding human secreted proteins are

XX KW used for the prevention, amelioration and treatment of autoimmune,

XX KW inflammatory, hyperproliferative and cardiovascular disorders, cancer,

XX KW PT wounds, and infectious diseases -

XX PS Disclosure; Page 152; 803pp; English.

XX CC The present invention relates to 12 secreted human proteins and the

XX CC nucleotide sequences encoding them. The polynucleotide sequences given

XX CC in AAA80606-A80623 encode the 12 secreted protein sequences given in

XX CC AAA25576-B25593. The human secreted proteins have various activities

XX CC dependent on the tissues in which they are expressed. Examples of the

XX CC activities of the proteins include: immunosuppressant;

XX CC anti-inflammatory; antiarthritic; antirheumatic, dermatological;

XX CC antiviral; antibacterial; and antifungal activity. The proteins,

XX CC and/or agonists and antagonists may be used to treat prevent

XX CC and/or diagnose various disease, disorders and conditions examples of

XX CC arthritis, dermatitis, and multiple sclerosis; inflammatory disorders

XX CC e.g. inflammatory bowel disease, Crohn's disease and nephritis;

XX CC hyperproliferative disorders such as paraproteinemia and purpura;

XX CC cardiovascular disorders e.g. coronary arteriosclerosis and myocarditis;

XX CC cancer e.g. melanoma and lymphoma. The proteins and polynucleotide

XX CC sequences may also be used in wound healing and the treatment of

XX CC infectious diseases. The human secreted protein gene #9 and protein

XX CC sequences are represented in sequences AAA80614 and AAB25584. Sequences

XX CC AAA80664-A80668 represent genes related to the secreted protein gene#9.

XX SQ Sequence 461 AA;

Query Match 100.0%; Score 26; DB 21; Length 461;

Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5

Db 285 KHAVE 289

|||||

RESULT 20

PT Isolated nucleic acids encoding human B7-like polypeptides, useful for
 PT diagnosis and treatment of e.g. inflammation, cancer, immune disorders
 PT such as Addison's disease, and cardiovascular disorders such as
 PT myocardial ischemias -

XX Example 1; Page 470-472; 493pp; English.

PS The present invention relates to novel human B7-like polypeptides and
 CC polynucleotides encoding such proteins. Sequences of the invention are
 CC used for preventing, treating or ameliorating a medical condition in a
 CC mammalian subject. The polynucleotides and polypeptides are administered
 CC to subjects having a disorder related to B-7 like polypeptides, such as
 CC inappropriate or excessive inflammation which can lead to tissue damage
 CC or even death, where the inflammation is brought about by the activation
 CC of certain cells in the body e.g. T cells and may involve disorders
 CC related to immune system. The nucleic acids, proteins, antibodies,
 CC agonists and antagonists of the invention are useful in the diagnosis,
 CC treatment and prevention of cancer (e.g. cancers of the adrenal gland,
 CC bone, bone marrow, breast, gastrointestinal tract, liver, urogenital or
 CC lung), immune disorders (e.g., Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis, ulcerative colitis),
 CC cardiovascular disorders (e.g., myocardial ischaemias), nervous system
 CC disorders (Alzheimer's disease, Parkinson's disease), endocrine disorders
 CC (e.g., diabetes mellitus, Grave's disease), reproductive system disorders
 CC (e.g., cryptorchism, Paget's disease), gastrointestinal disorders (e.g.,
 CC dysphagia, irritable bowel syndrome), liver disorders (e.g., hepatitis,
 CC hepatomegaly), neurological diseases (e.g., cerebral anoxia and epilepsy)
 CC and infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Sequences of the invention are also used in gene therapy.
 CC The present sequence is human B7-H9 protein.

SQ Sequence 461 AA;

Query Match 100.0%; Score 26; DB 23; Length 461;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5

Db 285 KHAVE 289

RESULT 22

AY99413
 ID AAY99413 standard; Protein; 466 AA.

XX AC AAY99413;

XX DT 08-AUG-2000 (first entry)

XX DE Human PRO1472 (UNQ744) amino acid sequence SEQ ID NO:267.

XX DE Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
 KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.

XX OS Homo sapiens.

XX PN WO200012708-A2.

XX PD 09-MAR-2000.

XX PF 01-SEP-1999; 99WO-US20111.

XX PR 01-SEP-1998; 98US-0098716.

XX PR 01-SEP-1998; 98US-0098749.

XX PR 01-SEP-1998; 98US-0098750.

XX PR 02-SEP-1998; 98US-0098803.

XX PR 02-SEP-1998; 98US-0098821.

XX PR 02-SEP-1998; 98US-0098843.

XX PR 09-SEP-1998; 98US-0099536.

XX PR 09-SEP-1998; 98US-0099596.

PR 09-SEP-1998; 98US-0099602.
 PR 09-SEP-1998; 98US-0099642.
 PR 10-SEP-1998; 98US-0099741.
 PR 10-SEP-1998; 98US-0099754.
 PR 10-SEP-1998; 98US-0099763.
 PR 10-SEP-1998; 98US-0099792.
 PR 10-SEP-1998; 98US-0099808.
 PR 10-SEP-1998; 98US-0099812.
 PR 10-SEP-1998; 98US-0099815.
 PR 10-SEP-1998; 98US-0099816.
 PR 15-SEP-1998; 98US-0100385.
 PR 15-SEP-1998; 98US-0100388.
 PR 15-SEP-1998; 98US-0100390.
 PR 16-SEP-1998; 98US-0100390.
 PR 16-SEP-1998; 98US-0100384.
 PR 16-SEP-1998; 98US-0100627.
 PR 16-SEP-1998; 98US-0100661.
 PR 16-SEP-1998; 98US-0100662.
 PR 16-SEP-1998; 98US-0100664.
 PR 17-SEP-1998; 98US-0100683.
 PR 17-SEP-1998; 98US-0100684.
 PR 17-SEP-1998; 98US-0100710.
 PR 17-SEP-1998; 98US-0100710.
 PR 17-SEP-1998; 98US-0100711.
 PR 17-SEP-1998; 98US-0100919.
 PR 18-SEP-1998; 98US-0100930.
 PR 18-SEP-1998; 98US-0100848.
 PR 18-SEP-1998; 98US-0100849.
 PR 18-SEP-1998; 98US-0101014.
 PR 18-SEP-1998; 98US-0101014.
 PR 18-SEP-1998; 98US-0101068.
 PR 18-SEP-1998; 98US-0101071.
 PR 23-SEP-1998; 98US-0101279.
 PR 23-SEP-1998; 98US-0101471.
 PR 23-SEP-1998; 98US-0101472.
 PR 23-SEP-1998; 98US-0101472.
 PR 23-SEP-1998; 98US-0101474.
 PR 23-SEP-1998; 98US-0101475.
 PR 23-SEP-1998; 98US-0101476.
 PR 23-SEP-1998; 98US-0101477.
 PR 23-SEP-1998; 98US-0101479.
 PR 24-SEP-1998; 98US-0101738.
 PR 24-SEP-1998; 98US-0101741.
 PR 24-SEP-1998; 98US-0101743.
 PR 24-SEP-1998; 98US-0101915.
 PR 24-SEP-1998; 98US-0101916.
 PR 24-SEP-1998; 98US-0102207.
 PR 24-SEP-1998; 98US-0102240.
 PR 29-SEP-1998; 98US-0102307.
 PR 29-SEP-1998; 98US-0102330.
 PR 29-SEP-1998; 98US-0102331.
 PR 30-SEP-1998; 98US-0102484.
 PR 30-SEP-1998; 98US-0102487.
 PR 30-SEP-1998; 98US-0102570.
 PR 30-SEP-1998; 98US-0102571.
 PR 01-OCT-1998; 98US-0102684.
 PR 01-OCT-1998; 98US-0102687.
 PR 02-OCT-1998; 98US-0102965.
 PR 06-OCT-1998; 98US-0103258.
 PR 06-OCT-1998; 98US-0103449.
 PR 07-OCT-1998; 98US-0103314.
 PR 07-OCT-1998; 98US-0103315.
 PR 07-OCT-1998; 98US-0103328.
 PR 07-OCT-1998; 98US-0103395.
 PR 07-OCT-1998; 98US-0103396.
 PR 07-OCT-1998; 98US-0103401.
 PR 08-OCT-1998; 98US-0103633.
 PR 08-OCT-1998; 98US-0103678.
 PR 08-OCT-1998; 98US-0103679.
 PR 14-OCT-1998; 98US-0103711.
 PR 14-OCT-1998; 98US-0104257.
 PR 20-OCT-1998; 98US-0104987.
 PR 20-OCT-1998; 98US-0105000.
 PR 21-OCT-1998; 98US-0105002.
 PR 21-OCT-1998; 98US-0105104.
 PR 22-OCT-1998; 98US-0105169.
 PR 22-OCT-1998; 98US-0105266.

PR 26-OCT-1998; 98US-0105693.
 PR 26-OCT-1998; 98US-0105694.
 PR 27-OCT-1998; 98US-0105807.
 PR 27-OCT-1998; 98US-0105881.
 PR 27-OCT-1998; 98US-0105882.
 PR 27-OCT-1998; 98US-0106062.
 PR 28-OCT-1998; 98US-0106023.
 PR 28-OCT-1998; 98US-0106029.
 PR 28-OCT-1998; 98US-0106030.
 PR 28-OCT-1998; 98US-0106032.
 PR 28-OCT-1998; 98US-0106033.
 PR 28-OCT-1998; 98US-0106178.
 PR 29-OCT-1998; 98US-0106248.
 PR 29-OCT-1998; 98US-0106384.
 PR 29-OCT-1998; 98US-0108500.
 PR 30-OCT-1998; 98US-0106464.
 PR 03-NOV-1998; 98US-0106856.
 PR 03-NOV-1998; 98US-0106902.
 PR 03-NOV-1998; 98US-0106905.
 PR 03-NOV-1998; 98US-0106919.
 PR 03-NOV-1998; 98US-0106932.
 PR 03-NOV-1998; 98US-0106934.
 PR 10-NOV-1998; 98US-0107783.
 PR 17-NOV-1998; 98US-0108775.
 PR 17-NOV-1998; 98US-0108779.
 PR 17-NOV-1998; 98US-0108787.
 PR 17-NOV-1998; 98US-0108788.
 PR 17-NOV-1998; 98US-0108801.
 PR 17-NOV-1998; 98US-0108802.
 PR 17-NOV-1998; 98US-0108806.
 PR 17-NOV-1998; 98US-0108807.
 PR 17-NOV-1998; 98US-0108867.
 PR 17-NOV-1998; 98US-0108925.
 PR 18-NOV-1998; 98US-0108848.
 PR 18-NOV-1998; 98US-0108849.
 PR 18-NOV-1998; 98US-0108850.
 PR 18-NOV-1998; 98US-0108851.
 PR 18-NOV-1998; 98US-0108852.
 PR 18-NOV-1998; 98US-0108858.
 PR 18-NOV-1998; 98US-0108904.

(GETH) GENENTECH INC.

Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;

WPI; 2000-237871/20.

N-PSDB; AAA37095.

PT New mammalian DNA sequences encoding transmembrane, receptor or
 PT secreted PRO polypeptides, useful for screening of potential peptide or
 PT small molecule inhibitors of the relevant receptor/ligand interactions

PS Claim 12; Fig 148; 773pp; English.

CC AAA37022 to AAA37144 encode the new isolated human transmembrane,
 CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
 CC transmembrane and receptor PRO proteins can be used for screening of
 CC potential peptide, or small molecule inhibitors of the relevant
 CC receptor/ligand interactions. The polypeptides and nucleotide sequences
 CC encoding then have various industrial applications, including uses as
 CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
 CC PCR primers and hybridisation probes used in the isolation of the PRO
 CC polypeptides from the present invention.

SQ Sequence 466 AA;

Query Match 100.0%; Score 26; DB 21; Length 466;

Best Local Similarity 100.0%; Pred. NO. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5

Db 285 KHAVE 289

RESULT 23

AAU29181

ID AAU29181 standard; Protein; 466 AA.

XX AC AAU29181;

XX DT 18-DEC-2001 (first entry)

XX DE Human PRO polypeptide sequence #158.

XX KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX OS Homo sapiens.

XX PN WO200168848-A2.

XX PD 20-SEP-2001.

XX PF 28-FEB-2001; 2001WO-US06520.

XX PR 01-MAR-2000; 2000WO-US05601.

XX PR 02-MAR-2000; 2000WO-US05841.

XX PR 03-MAR-2000; 2000US-187202P.

XX PR 06-MAR-2000; 2000US-186968P.

XX PR 14-MAR-2000; 2000US-189320P.

XX PR 14-MAR-2000; 2000US-189328P.

XX PR 15-MAR-2000; 2000WO-US06884.

XX PR 21-MAR-2000; 2000US-190828P.

XX PR 21-MAR-2000; 2000US-191007P.

XX PR 21-MAR-2000; 2000US-191048P.

XX PR 21-MAR-2000; 2000US-191314P.

XX PR 28-MAR-2000; 2000US-192855P.

XX PR 29-MAR-2000; 2000US-193032P.

XX PR 30-MAR-2000; 2000US-193053P.

XX PR 04-APR-2000; 2000US-194449P.

XX PR 04-APR-2000; 2000US-194647P.

XX PR 11-APR-2000; 2000US-195975P.

XX PR 11-APR-2000; 2000US-196000P.

XX PR 11-APR-2000; 2000US-196187P.

XX PR 11-APR-2000; 2000US-196690P.

XX PR 11-APR-2000; 2000US-196820P.

XX PR 18-APR-2000; 2000US-198121P.

XX PR 18-APR-2000; 2000US-198585P.

XX PR 25-APR-2000; 2000US-199397P.

XX PR 25-APR-2000; 2000US-199550P.

XX PR 25-APR-2000; 2000US-199654P.

XX PR 03-MAY-2000; 2000US-201516P.

XX PR 17-MAY-2000; 2000WO-US13705.

XX PR 22-MAY-2000; 2000WO-US14042.

XX PR 30-MAY-2000; 2000WO-US14941.

XX PR 02-JUN-2000; 2000WO-US15264.

XX PR 05-JUN-2000; 2000US-209832P.

XX PR 28-JUL-2000; 2000WO-US20710.

XX PR 22-AUG-2000; 2000US-0644848.

XX PR 24-AUG-2000; 2000WO-US23328.

XX PR 08-NOV-2000; 2000WO-US30952.

XX PR 01-DEC-2000; 2000WO-US32678.

XX PR 20-DEC-2000; 2000WO-US34956.

XX (GETH) GENENTECH INC.

PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;

XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

DR WPI; 2001-602746/68.

DR N-PSDB; AAS46082.

PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and
 PT to screen for modulators of the compounds -
 XX
 XX Claim 11; Fig 316; 774pp; English.

XX Sequences AAU29024-AAU29128 represent PRO polypeptides of the invention.
 CC The PRO polypeptides and their associated nucleic acids can be used to
 CC detect the presence of a tumour in a mammal by comparing the level of
 CC expression of a PRO polypeptide in a test sample of cells from the animal
 CC and a control sample of normal cells, whereby a higher level of
 CC expression in the test sample indicates the presence of a tumour in the
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
 CC when contacted with it. A specific polypeptide can be used to stimulate
 CC the proliferation, or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence: 466 AA;

Query Match 100.0%; Score 26; DB 22; Length 466;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
 |||||
 Db 285 KHAVE 289

RESULT 24

AA874731
 ID AAB74731 standard; Protein; 466 AA.

XX
 AC AAB74731;

XX 12-JUN-2001 (first entry)

XX Human membrane associated protein MEMAP-37.

XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
 KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
 KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
 KW autoimmune disorder; inflammatory disorder; neurological disorder;
 KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
 KW epilepsy; diarrhoea.

XX Homo sapiens.

XX WO200112662-A2.

XX 22-FEB-2001.

XX 14-AUG-2000; 2000WO-US22315.

XX 17-AUG-1999; 9SUS-0149641.

XX 09-NOV-1999; 9SUS-0164203.

XX (INCY-) INCYTE GENOMICS INC.

XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;

PI Baughn MR, Lu DAM, Patterson C;

XX WPI; 2001-168860/17.

DR N-PSDB; AAF81777.

XX Isolated polypeptide with a human membrane associated protein sequence
 PT is useful for the diagnosis, prevention and treatment of cell
 PT proliferative, autoimmune/inflammatory, neurological and

PT gastrointestinal disorders -

XX Claim 1; Page 148-149; 173pp; English.

XX AAF81741 to AAF81777 encode the human membrane associated proteins
 CC (MEMAP), given in AAB74695 to AAB74731. MEMAPs have cyostatic,
 CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and
 CC antiarteriosclerotic activities, which can be used in gene therapy.
 CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
 CC associated with decreased expression of functional MEMAP and antagonists
 CC of MEMAP are used to treat a disease or condition associated with
 CC overexpression of functional MEMAP. These disorders include cell
 CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal
 CC disorders. The MEMAP polynucleotides and proteins are also used for the
 CC diagnosis of these disorders. Specific examples of these disorders
 CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
 CC MEMAP proteins can be used to screen for compounds which specifically
 CC bind MEMAP including antibodies, oligonucleotides, proteins and small
 CC molecules. MEMAP polynucleotides can be used to prepare transgenic
 CC animals which can be studied to provide information concerning human
 CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
 CC detection of MEMAP protein and can be used as antagonists for the
 CC prevention of a disorder associated with MEMAP. Polynucleotides encoding MEMAP
 CC can be delivered to target cells with genetic abnormalities with respect
 CC to the expression of MEMAP to treat or prevent a disorder associated
 CC with MEMAP.

XX Sequence 466 AA;

Query Match 100.0%; Score 26; DB 22; Length 466;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
 |||||
 Db 285 KHAVE 289

RESULT 25

AA87577

ID AAB87577 standard; Protein; 466 AA.

XX
 AC AAB87577;

XX 15-MAY-2001 (first entry)

XX Human PRO1472.

DE Human PRO1472.

XX Human; PRO protein; mapping.

XX Homo sapiens.

XX WO200116318-A2.

XX 08-MAR-2001.

XX 24-AUG-2000; 2000WO-US23328.

XX 01-SEP-1999; 99WO-US20111.

XX 15-SEP-1999; 99WO-US21090.

XX 07-DEC-1999; 99US-0169495.

XX 09-DEC-1999; 99US-0170262.

XX 11-JAN-2000; 2000US-0175481.

XX 18-FEB-2000; 2000WO-US04341.

XX 22-FEB-2000; 2000WO-US04342.

XX 01-MAR-2000; 2000WO-US04414.

XX 01-MAR-2000; 2000WO-US05601.

XX 03-MAR-2000; 2000US-0187202.

XX 25-APR-2000; 2000US-0199397.

XX 22-MAY-2000; 2000WO-US14042.

XX 05-JUN-2000; 2000US-0209832.

XX (GETH) GENENTECH INC.

XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
DR N-PSDB; AAF92109.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosome and gene mapping.
XX
PS Claim 12; Fig 104; 278pp; English.
XX
CC The present sequence is a human PRO polypeptide (secreted and
CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC anti-PRO antibodies are useful for preparation of a medicament useful in
CC the treatment of a condition which is responsive to the PRO protein,
CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC employed as molecular weight markers for protein electrophoresis. The PRO
CC coding sequence has applications in molecular biology, including use as
CC hybridisation probes, and in chromosome and gene mapping.
XX
SQ Sequence 466 AA;
Query Match 100.0%; Score 26; DB 22; Length 466;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
Db 285 KHAVE 289
|||||

Search completed: June 20, 2003, 20:51:57
Job time : 33.9318 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 10.5682 Seconds

(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-16

Perfect score: 26

Sequence: 1 KHAVE 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/6C COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	5	4	US-08-996-679-56
2	26	100.0	5	4	US-09-115-395-18
3	26	100.0	5	4	US-09-250-059-16
4	26	100.0	5	4	US-09-248-074-16
5	26	100.0	5	4	US-09-357-717-14
6	26	100.0	5	4	US-09-458-870-16
7	26	100.0	266	4	US-09-631-828A-4
8	26	100.0	588	3	US-09-193-107-2
9	26	100.0	588	4	US-09-477-472-2
10	26	100.0	1048	4	US-09-171-699-10
11	25	96.2	139	1	US-08-680-726A-66
12	25	96.2	139	4	US-09-092-409-66
13	25	96.2	425	4	US-09-293-549-6
14	25	96.2	665	4	US-09-382-106-2
15	25	96.2	724	4	US-08-914-999-4
16	25	96.2	1038	4	US-09-541-782-4
17	25	96.2	1038	4	US-09-723-820-4
18	23	88.5	5	3	US-08-893-534A-20
19	23	88.5	5	4	US-08-996-679-20
20	23	88.5	5	4	US-09-115-395-53
21	23	88.5	5	4	US-09-507-102-20
22	23	88.5	5	4	US-09-250-059-12
23	23	88.5	5	4	US-09-248-074-12
24	23	88.5	5	4	US-09-357-717-12
25	23	88.5	5	4	US-09-458-870-12
26	23	88.5	27	3	US-09-045-632-64
27	23	88.5	84	3	US-09-045-632-8

28	88.5	117	4	US-09-149-476-642	Sequence 642, Appl
29	88.5	182	3	US-09-045-632-29	Sequence 29, Appl
30	88.5	185	3	US-09-045-632-26	Sequence 26, Appl
31	88.5	220	1	US-08-233-146-2	Sequence 2, Appl
32	88.5	220	1	US-08-463-470-2	Sequence 2, Appl
33	88.5	222	1	US-07-732-242C-7	Sequence 7, Appl
34	88.5	255	4	US-09-134-001C-3680	Sequence 3680, Ap
35	88.5	283	3	US-09-045-632-27	Sequence 27, Appl
36	88.5	327	6	5171684-6	Patent No. 5171684
37	88.5	339	4	US-09-149-476-488	Sequence 488, App
38	88.5	357	4	US-09-149-476-755	Sequence 755, App
39	88.5	404	3	US-09-045-632-23	Sequence 23, Appl
40	88.5	462	2	US-08-477-451-24	Sequence 24, Appl
41	88.5	478	2	US-08-951-148-1	Sequence 1, Appl
42	88.5	478	2	US-08-951-148-7	Sequence 7, Appl
43	88.5	478	2	US-09-165-234-1	Sequence 1, Appl
44	88.5	478	2	US-09-165-234-7	Sequence 7, Appl
45	88.5	478	3	US-09-274-570-1	Sequence 1, Appl
46	88.5	478	3	US-09-274-570-7	Sequence 7, Appl
47	88.5	478	4	US-09-440-936-4	Sequence 4, Appl
48	88.5	498	3	US-09-045-632-30	Sequence 30, Appl
49	88.5	502	3	US-09-045-632-24	Sequence 24, Appl
50	88.5	504	3	US-09-045-632-19	Sequence 19, Appl
51	88.5	541	3	US-09-045-632-36	Sequence 36, Appl
52	88.5	599	3	US-09-045-632-28	Sequence 28, Appl
53	88.5	602	3	US-09-045-632-20	Sequence 20, Appl
54	88.5	604	3	US-09-045-632-14	Sequence 14, Appl
55	88.5	605	4	US-09-440-936-2	Sequence 2, Appl
56	88.5	642	3	US-09-045-632-35	Sequence 35, Appl
57	88.5	667	4	US-09-342-647-28	Sequence 28, Appl
58	88.5	702	3	US-09-045-632-15	Sequence 15, Appl
59	88.5	818	3	US-09-045-632-25	Sequence 25, Appl
60	88.5	861	3	US-09-045-632-34	Sequence 34, Appl
61	88.5	918	3	US-09-045-632-21	Sequence 21, Appl
62	88.5	961	3	US-09-045-632-33	Sequence 33, Appl
63	88.5	1018	3	US-09-045-632-16	Sequence 16, Appl
64	88.5	1061	3	US-09-045-632-32	Sequence 32, Appl
65	88.5	1112	3	US-09-045-632-2	Sequence 2, Appl
66	88.5	1112	3	US-09-045-632-3	Sequence 3, Appl
67	88.5	2475	4	US-09-413-814-48	Sequence 48, Appl
68	84.6	5	4	US-09-115-395-73	Sequence 73, Appl
69	84.6	5	4	US-09-250-059-17	Sequence 17, Appl
70	84.6	5	4	US-09-248-074-17	Sequence 17, Appl
71	84.6	5	4	US-09-458-870-17	Sequence 17, Appl
72	84.6	40	3	US-08-258-287B-48	Sequence 48, Appl
73	84.6	40	3	US-08-368-704C-47	Sequence 47, Appl
74	84.6	45	2	US-08-483-926A-2	Sequence 2, Appl
75	84.6	45	2	US-08-854-768-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-08-996-679-56
; Sequence 56, Application US/08996679
; Patent No. 6169071
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; NUMBER OF INVENTION: CELL ADHESION
; CORRESPONDENCE ADDRESS: 63
; ADDRESS: SEED AND BERRY LLP
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996.679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-56

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVE 5
Db 1 KHAVE 5

RESULT 2
US-09-115-395-18
Sequence 18, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115.395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996.679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893.534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021.612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 5
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis
US-09-115-395-18

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVE 5
Db 1 KHAVE 5

RESULT 3
US-09-250-059-16
Sequence 16, Application US/09250059
Patent No. 6333307
GENERAL INFORMATION:

APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250.059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 5
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-250-059-16

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVE 5
Db 1 KHAVE 5

RESULT 4
US-09-248-074-16
Sequence 16, Application US/09248074
Patent No. 6346512
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
FILE REFERENCE: 100086.401C5
CURRENT APPLICATION NUMBER: US/09/248.074
CURRENT FILING DATE: 1999-02-10
NUMBER OF SEQ ID NOS: 81
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 5
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-248-074-16

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVE 5
Db 1 KHAVE 5

RESULT 5
US-09-357-717-14
Sequence 14, Application US/09357717
Patent No. 6417325

GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farooqui, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 10086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-357-717-14

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db |||||
1 KHAVE 5

RESULT 6
US-09-458-870-16
; Sequence 16, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farooqui, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 10086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-16

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db |||||
1 KHAVE 5

RESULT 7
US-09-631-828A-4
; Sequence 4, Application US/09631828A
; Patent No. 6461852
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, NOBUHARU
; APPLICANT: YASUEDA, HISASHI
; APPLICANT: KAWAHARA, YOSHIO
; APPLICANT: SUGIMOTO, SHINICHI
; TITLE OF INVENTION: GENES FOR LYSINE BIOSYNTHETIC SYSTEM DERIVED FROM THERMOPHIL
; FILE REFERENCE: 195277US0
; CURRENT APPLICATION NUMBER: US/09/631,828A
; CURRENT FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 11-221468
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Bacillus methanolicus
US-09-631-828A-4

Query Match 100.0%; Score 26; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db |||||
245 KHAVE 249

RESULT 8
US-09-193-107-2
; Sequence 2, Application US/09193107
; Patent No. 6033892
; GENERAL INFORMATION:
; APPLICANT: Gregory Gambetta
; TITLE OF INVENTION: Polypeptides Having 5-Aminolevulinic
; TITLE OF INVENTION: Acid Synthases And Nucleic Acids Encoding Same
; FILE REFERENCE: 5399.200-US
; CURRENT APPLICATION NUMBER: US/09/193,107
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Fusarium
US-09-193-107-2

Query Match 100.0%; Score 26; DB 3; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db |||||
198 KHAVE 202

RESULT 9
US-09-477-472-2
; Sequence 2, Application US/09477472
; Patent No. 6207433
; GENERAL INFORMATION:
; APPLICANT: Gregory Gambetta
; TITLE OF INVENTION: Polypeptides Having 5-Aminolevulinic
; TITLE OF INVENTION: Acid Synthase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5399.210-US
; CURRENT APPLICATION NUMBER: US/09/477,472
; CURRENT FILING DATE: 2000-01-04
; EARLIER APPLICATION NUMBER: 09/193,107
; EARLIER FILING DATE: 1998-11-16

;; EARLIER APPLICATION NUMBER: 60/066,107
;; EARLIER FILING DATE: 1997-11-17
;; NUMBER OF SEQ ID NOS: 12
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2
;; TYPE: PRT
;; LENGTH: 588
;; ORGANISM: Fusarium
US-09-477-472-2

Query Match 100.0%; Score 26; DB 4; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 198 KHAVE 202

RESULT 10
US-09-171-699-10
; Sequence 10, Application US/09171699
; Patent No. 6448389
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Genczol, Eva
; Berencsi, Klara
; Kari, Csaba
; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and
; Uses Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/171,699
FILING DATE: 19-Jan-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/015,717
FILING DATE: 23-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kodroff, Cathy A.
REGISTRATION NUMBER: 33,980
REFERENCE/DOCKET NUMBER: WST66APCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1048 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-171-699-10

Query Match 100.0%; Score 26; DB 4; Length 1048;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 255 KHAVE 259

RESULT 11
US-08-680-726A-66
; Sequence 66, Application US/08680726A
; Patent No. 5804197
; GENERAL INFORMATION:
; APPLICANT: Haanes, Elizabeth J.
; APPLICANT: Frank, Rexann S.
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,726A
; FILING DATE: 12-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-46-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-680-726A-66

Query Match 96.2%; Score 25; DB 1; Length 139;
Best Local Similarity: 80.0%; Pred. No. 48;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 53 KHAIE 57

RESULT 12
US-09-092-409-66
; Sequence 66, Application US/09092409
; Patent No. 6159478
; GENERAL INFORMATION:
; APPLICANT: Haanes, Elizabeth J.
; APPLICANT: Frank, Rexann S.
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/092,409
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/680,726
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-092-409-66

Query Match 96.2%; Score 25; DB 4; Length 139;
Best Local Similarity 80.0%; Pred. No. 48;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
DB 53 KHAIE 57

RESULT 13
US-09-293-549-6
Sequence 6, Application US/09293549
Patent No. 644049
GENERAL INFORMATION:
APPLICANT: G. Todd Milne
APPLICANT: Gerald Fink
TITLE OF INVENTION: A METHOD FOR EFFICIENT AND HIGHLY
TITLE OF INVENTION: SELECTIVE CONTROL OF MICROORGANISMS
FILE REFERENCE: 50078/008002
CURRENT APPLICATION NUMBER: US/09/293,549
CURRENT FILING DATE: 1999-04-16
EARLIER APPLICATION NUMBER: 60/082,089
EARLIER FILING DATE: 1998-04-17
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 425
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-293-549-6

Query Match 96.2%; Score 25; DB 4; Length 425;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
DB 421 KHAIE 425

RESULT 14
US-09-382-106-2
Sequence 2, Application US/09382106
Patent No. 6221631
GENERAL INFORMATION:
APPLICANT: Huang, Jianzhong
APPLICANT: Jiang, Xinhe
APPLICANT: McDavitt, Damien
APPLICANT: Van Horn, Stephanie
TITLE OF INVENTION: tkTA
FILE REFERENCE: GM10236
CURRENT APPLICATION NUMBER: US/09/382,106

CURRENT FILING DATE: 1999-08-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-382-106-2

Query Match 96.2%; Score 25; DB 4; Length 665;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
DB 504 KHAIE 508

RESULT 15
US-08-914-999-4
Sequence 4, Application US/08914999
Patent No. 6346406
GENERAL INFORMATION:
APPLICANT: Ryazanov, Alexey G.
APPLICANT: Hait, William N.
APPLICANT: Pavur, Karen S.
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
TITLE OF INVENTION: AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-914-999-4

Query Match 96.2%; Score 25; DB 4; Length 724;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
DB 85 KHAIE 89

RESULT 16

US-09-541-782-4
; Sequence 4, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541.782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1038
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-541-782-4

Query Match 96.2%; Score 25; DB 4; Length 1038;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 1013 KHAIE 1017

RESULT 17

US-09-723-820-4
; Sequence 4, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723.820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1038
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-723-820-4

Query Match 96.2%; Score 25; DB 4; Length 1038;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 1013 KHAIE 1017

RESULT 18

US-08-893-534A-20
; Sequence 20, Application US/08893534A
; Patent No. 6031072
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,534A
FILING DATE: 11-JUL-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-893-534A-20

Query Match 88.5%; Score 23; DB 3; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 1 KHAVD 5

RESULT 19

US-08-996-679-20
; Sequence 20, Application US/08996679
; Patent No. 6169071
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-20

Query Match 88.5%; Score 23; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 1 KHAVD 5

RESULT 20

US-09-115-395-53
Sequence 53, Application US/09115395A
Patent No. 6207639

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115,395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996,679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893,534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 53
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis

US-09-115-395-53

Query Match 88.5%; Score 23; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 1 KHAVD 5

RESULT 21

US-09-507-102-20
Sequence 20, Application US/09507102
Patent No. 6326352

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED IP LAW GROUP PLLC
STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,102
FILING DATE: 17-Feb-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/893,534
FILING DATE: 11-JUL-1997
APPLICATION NUMBER: US 60/021,612
FILING DATE: 12-JUL-1996

ATTORNEY/AGENT INFORMATION:
NAME: Christiansen, William T.
REGISTRATION NUMBER: 44,614
REFERENCE/DOCKET NUMBER: 100086.401C10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-507-102-20

Query Match 88.5%; Score 23; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 1 KHAVD 5

RESULT 22

US-09-250-059-12
Sequence 12, Application US/09250059
Patent No. 633307

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250,059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: Peptide with cadherin cell adhesion recognition
OTHER INFORMATION: sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-250-059-12

Query Match 88.5%; Score 23; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 1 KHAVD 5

RESULT 23
US-09-248-074-12
; Sequence 12, Application US/09248074
; Patent No. 6346512
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C5
; CURRENT APPLICATION NUMBER: US/09/248,074
; CURRENT FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with cadherin cell adhesion recognition
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-248-074-12

Query Match 88.5%; Score 23; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
Db 1 KHAWE 5

RESULT 24
US-09-357-717-12
; Sequence 12, Application US/09357717
; Patent No. 6417325
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with cadherin cell adhesion recognition
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-357-717-12

Query Match 88.5%; Score 23; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5

Db 1 KHAWE 5
RESULT 25
US-09-458-870-12
; Sequence 12, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with cadherin cell adhesion recognition
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-12

Query Match 88.5%; Score 23; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
Db 1 KHAWE 5

Search completed: June 20, 2003, 21:00:31
Job time : 11.5682 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:52:19 ; Search time 12.8409 Seconds
(without alignments)
42.134 Million cell updates/sec

Title: US-10-105-008-16

Perfect score: 26

Sequence: 1 KHAVE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pap:
2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pap:
3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pap:
4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.pap:
5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pap:
6: /cgn2_6/ptodata/2/pubpaa/US07 PUBCOMB.pap:
7: /cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pap:
8: /cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB.pap:
9: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pap:
10: /cgn2_6/ptodata/2/pubpaa/US09 PUBCOMB.pap:
11: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pap:
12: /cgn2_6/ptodata/2/pubpaa/US10 PUBCOMB.pap:
13: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pap:
14: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	26	100.0	5	US-09-769-145-16
2	26	100.0	5	US-10-058-821-14
3	26	100.0	5	US-09-234-328-311
4	26	100.0	5	US-09-305-928-311
5	26	100.0	266	US-10-214-556-4
6	26	100.0	378	US-09-816-664-7
7	26	100.0	415	US-09-800-198-89
8	26	100.0	461	US-09-984-130-44
9	26	100.0	461	US-09-984-130-119
10	26	100.0	466	US-09-965-529-37
11	26	100.0	466	US-10-063-547-104
12	26	100.0	466	US-10-174-590-316
13	26	100.0	466	US-10-176-758-316
14	26	100.0	466	US-10-063-616-104
15	26	100.0	466	US-10-175-737-316
16	26	100.0	466	US-10-063-502-104
17	26	100.0	466	US-10-173-706-316
18	26	100.0	466	US-10-175-738-316
19	26	100.0	466	US-10-175-752-316

ALIGNMENTS

RESULT 1:
US-09-769-145-16

; Sequence 16, Application US/09769145
; Patent No. US20020168761A1

; GENERAL INFORMATION:

; APPLICANT: Gour, Barbara J.

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Ali, Ammar

; APPLICANT: Ni, Feng

; APPLICANT: Chen, Zhigang

; APPLICANT: Michaud, Stephanie

; APPLICANT: Wang, Shoaming

20	26	100.0	466	9	US-10-176-482-316	Sequence 316, App
21	26	100.0	466	9	US-10-176-757-316	Sequence 316, App
22	26	100.0	466	9	US-10-176-913-316	Sequence 316, App
23	26	100.0	466	9	US-10-180-552-316	Sequence 316, App
24	26	100.0	466	9	US-10-180-557-316	Sequence 316, App
25	26	100.0	466	9	US-10-173-700-316	Sequence 316, App
26	26	100.0	466	9	US-10-174-572-316	Sequence 316, App
27	26	100.0	466	9	US-10-174-579-316	Sequence 316, App
28	26	100.0	466	9	US-10-174-582-316	Sequence 316, App
29	26	100.0	466	9	US-10-174-588-316	Sequence 316, App
30	26	100.0	466	9	US-10-175-739-316	Sequence 316, App
31	26	100.0	466	9	US-10-175-740-316	Sequence 316, App
32	26	100.0	466	9	US-10-175-743-316	Sequence 316, App
33	26	100.0	466	9	US-10-176-488-316	Sequence 316, App
34	26	100.0	466	9	US-10-176-492-316	Sequence 316, App
35	26	100.0	466	9	US-10-176-747-316	Sequence 316, App
36	26	100.0	466	9	US-10-176-750-316	Sequence 316, App
37	26	100.0	466	9	US-10-176-985-316	Sequence 316, App
38	26	100.0	466	9	US-10-176-987-316	Sequence 316, App
39	26	100.0	466	9	US-10-176-991-316	Sequence 316, App
40	26	100.0	466	9	US-10-176-992-316	Sequence 316, App
41	26	100.0	466	9	US-10-176-993-316	Sequence 316, App
42	26	100.0	466	9	US-10-184-658-316	Sequence 316, App
43	26	100.0	466	9	US-10-173-695-316	Sequence 316, App
44	26	100.0	466	9	US-10-173-693-316	Sequence 316, App
45	26	100.0	466	9	US-10-173-705-316	Sequence 316, App
46	26	100.0	466	9	US-10-174-576-316	Sequence 316, App
47	26	100.0	466	9	US-10-174-585-316	Sequence 316, App
48	26	100.0	466	9	US-10-175-747-316	Sequence 316, App
49	26	100.0	466	9	US-10-176-481-316	Sequence 316, App
50	26	100.0	466	9	US-10-176-485-316	Sequence 316, App
51	26	100.0	466	9	US-10-176-483-316	Sequence 316, App
52	26	100.0	466	9	US-10-176-487-316	Sequence 316, App
53	26	100.0	466	9	US-10-176-493-316	Sequence 316, App
54	26	100.0	466	9	US-10-176-755-316	Sequence 316, App
55	26	100.0	466	9	US-10-176-911-316	Sequence 316, App
56	26	100.0	466	9	US-10-176-913-316	Sequence 316, App
57	26	100.0	466	9	US-10-176-925-316	Sequence 316, App
58	26	100.0	466	9	US-10-176-978-316	Sequence 316, App
59	26	100.0	466	9	US-10-179-510-316	Sequence 316, App
60	26	100.0	466	9	US-10-180-543-316	Sequence 316, App
61	26	100.0	466	9	US-10-180-544-316	Sequence 316, App
62	26	100.0	466	9	US-10-180-546-316	Sequence 316, App
63	26	100.0	466	9	US-10-180-547-316	Sequence 316, App
64	26	100.0	466	9	US-10-180-549-316	Sequence 316, App
65	26	100.0	466	9	US-10-180-553-316	Sequence 316, App
66	26	100.0	466	9	US-10-180-553-316	Sequence 316, App
67	26	100.0	466	9	US-10-181-000-316	Sequence 316, App
68	26	100.0	466	9	US-10-183-010-316	Sequence 316, App
69	26	100.0	466	9	US-10-183-012-316	Sequence 316, App
70	26	100.0	466	9	US-10-184-614-316	Sequence 316, App
71	26	100.0	466	9	US-10-184-623-316	Sequence 316, App
72	26	100.0	466	9	US-10-184-635-316	Sequence 316, App
73	26	100.0	466	9	US-10-184-637-316	Sequence 316, App
74	26	100.0	466	9	US-10-184-646-316	Sequence 316, App
75	26	100.0	466	9	US-10-184-647-316	Sequence 316, App

APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US/09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-16

Query Match 100.0%; Score 26; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
|
|
|
|
Db 1 KHAVE 5

RESULT 2

US-10-058-821-14
Sequence 14, Application US/10058821
Publication No. US20030087811A1

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
FILE REFERENCE: 100086.401C12
CURRENT APPLICATION NUMBER: US/10/058,821
CURRENT FILING DATE: 2002-01-29
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-10-058-821-14

Query Match 100.0%; Score 26; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
|
|
|
|
Db 1 KHAVE 5

RESULT 3

US-09-234-395-311
Sequence 311, Application US/09234395
Patent No. US20020123044A1

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 311
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-234-395-311

Query Match 100.0%; Score 26; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
|
|
|
|
Db 1 KHAVE 5

RESULT 4

US-09-305-928-311
Sequence 311, Application US/09305928
Patent No. US20020146687A1

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C4
CURRENT APPLICATION NUMBER: US/09/305,928
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 311
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-305-928-311

Query Match 100.0%; Score 26; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
|
|
|
|
Db 1 KHAVE 5

RESULT 5

US-10-214-556-4
Sequence 4, Application US/10214556
Publication No. US20030013174A1

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, NOBUHARU
APPLICANT: YASUEDA, HISASHI

```
; APPLICANT: KAWAHARA, YOSHIO
; APPLICANT: SUGIMOTO, SHINICHI
; TITLE OF INVENTION: GENES FOR LYSINE BIOSYNTHETIC SYSTEM DERIVED FROM THERMOPHILIC B
; FILE REFERENCE: 195277US0
; CURRENT APPLICATION NUMBER: US/10/214,556
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/09/631,828
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 11-221468
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Bacillus methanolicus
US-10-214-556-4

Query Match      100.0%; Score 26; DB 9; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHAWE 5
Db      245 KHAWE 249

RESULT 6
US-09-816-664-7
; Sequence 7, Application US/09816664
; Patent No. US20020038014A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 26443 and 46873, NOVEL HUMAN
; FILE REFERENCE: 10448-029001
; CURRENT APPLICATION NUMBER: US/09/816,664
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,973
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-816-664-7

Query Match      100.0%; Score 26; DB 10; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHAWE 5
Db      39 KHAWE 43

RESULT 7
US-09-800-198-89
; Sequence 89, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; APPLICANT: KAWAHARA, YOSHIO
; APPLICANT: SUGIMOTO, SHINICHI
; TITLE OF INVENTION: GENES FOR LYSINE BIOSYNTHETIC SYSTEM DERIVED FROM THERMOPHILIC B
; FILE REFERENCE: 195277US0
; CURRENT APPLICATION NUMBER: US/10/214,556
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/09/631,828
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 11-221468
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Halocynthia roretzi
US-09-800-198-89

Query Match      100.0%; Score 26; DB 9; Length 415;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHAWE 5
Db      202 KHAWE 206

RESULT 8
US-09-984-130-44
; Sequence 44, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (234)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (236)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-984-130-44

Query Match      100.0%; Score 26; DB 9; Length 461;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHAWE 5
Db      285 KHAWE 289

RESULT 9
US-09-984-130-119
; Sequence 119, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
```

FILE REFERENCE: PF489P2
CURRENT APPLICATION NUMBER: US/09/984,130
CURRENT FILING DATE: 2001-10-29
PRIOR FILING DATE: 2000-10-30
PRIOR APPLICATION NUMBER: 60/243,792
PRIOR FILING DATE: 2000-10-30
PRIOR APPLICATION NUMBER: 09/836,353
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/198,407
PRIOR FILING DATE: 2000-04-19
PRIOR APPLICATION NUMBER: PCT/US99/25031
PRIOR FILING DATE: 1999-10-27
PRIOR APPLICATION NUMBER: 60/105,971
PRIOR FILING DATE: 1998-10-28
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 119
LENGTH: 461
TYPE: PRT
ORGANISM: Homo sapiens
US-09-984-130-119

Query Match 100.0%; Score 26; DB 9; Length 461;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 10

US-09-965-529-37
Sequence 37, Application US/09965529

Publication No. US20020182671A1

GENERAL INFORMATION:

APPLICANT: LAL, Preeti
APPLICANT: YUE, Henry
APPLICANT: TANG, Y. Tom
APPLICANT: BANDMAN, Olga
APPLICANT: BURFORD, Neil
APPLICANT: AZIMZAI, Yalda
APPLICANT: BAUGHN, Maria R.
APPLICANT: LU, Dying Aina M.
APPLICANT: PATTERSON, Chandra
TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS

FILE REFERENCE: PF-0731 USA

CURRENT APPLICATION NUMBER: US/09/965,529

CURRENT FILING DATE: 2001-09-26

PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14

NUMBER OF SEQ ID NOS: 74

SOFTWARE: PERL Program

SEQ ID NO 37

LENGTH: 466

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No. US20020182671A1 6114480CD1

US-09-965-529-37

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 11

US-10-063-547-104

Sequence 104, Application US/10063547

Publication No. US20020182638A1

GENERAL INFORMATION:

APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3230R1C1

CURRENT APPLICATION NUMBER: US/10/063,547

CURRENT FILING DATE: 2002-05-02

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 170

SEQ ID NO 104

LENGTH: 466

TYPE: PRT

ORGANISM: Homo Sapien

US-10-063-547-104

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 12

US-10-174-590-316

Sequence 316, Application US/10174590

Publication No. US20030008352A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C42

CURRENT APPLICATION NUMBER: US/10/174,590

CURRENT FILING DATE: 2002-06-18

Prior application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 316

LENGTH: 466

TYPE: PRT

ORGANISM: Homo Sapien

US-10-174-590-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 13

US-10-176-758-316

; Sequence 316, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 14
US-10-063-616-104
; Sequence 104, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 104
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-616-104

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 15

US-10-175-737-316
; Sequence 316, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 16
US-10-063-502-104
; Sequence 104, Application US/10063502
; Publication No. US20030023042A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,502
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 104
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-502-104

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
Db 285 KHAVE 289

RESULT 17

US-10-173-706-316
; Sequence 316, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
Db 285 KHAWE 289

RESULT 18

US-10-175-738-316
; Sequence 316, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
Db 285 KHAWE 289

Db

285 KHAWE 289

RESULT 19

US-10-175-752-316
; Sequence 316, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-752-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
Db 285 KHAWE 289

RESULT 20

US-10-176-482-316
; Sequence 316, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAIVE 5
|||||
Db 285 KHAIVE 289

RESULT 21

US-10-176-757-316
; Sequence 316, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAIVE 5
|||||
Db 285 KHAIVE 289

RESULT 22

US-10-176-913-316
; Sequence 316, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-913-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAIVE 5
|||||
Db 285 KHAIVE 289

RESULT 23

US-10-180-552-316
; Sequence 316, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C53
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAIVE 5
|||||
Db 285 KHAIVE 289

RESULT 24

US-10-180-557-316
; Sequence 316, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
US-10-180-557-316

QY 1 KHAIVE 5
|||||
Db 285 KHAIVE 289

; ORGANISM: Homo Sapien
US-10-180-557-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 285 KHAVE 289

RESULT 25

US-10-173-700-316
; Sequence 316, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deenovers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabé, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P1430R1C14
; CURRENT APPLICATION NUMBER: US/10/173,700
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 316
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-700-316

Query Match 100.0%; Score 26; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 285 KHAVE 289

Search completed: June 20, 2003, 21:02:35
Job time : 14.8409 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 12.7273 Seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-16

Perfect score: 26

Sequence: 1 KUAVE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 75 summaries

Database :

PIR-73:***
1: PIR1:***
2: PIR2:***
3: PIR3:***
4: PIR4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	73	2 G87164	hypothetical prote
2	26	100.0	153	2 C69193	ribosomal protein
3	26	100.0	155	2 G69489	LSU ribosomal prot
4	26	100.0	267	2 H96668	protein F1N19.9 [i
5	26	100.0	268	2 A82604	oxidoreductase XP2
6	26	100.0	305	2 T16465	hypothetical prote
7	26	100.0	340	2 A97347	oligopeptide ABC t
8	26	100.0	349	2 B82487	hypothetical prote
9	26	100.0	417	2 T39939	DNA binding protel
10	26	100.0	418	2 AC3519	isochoismate synt
11	26	100.0	423	1 KGRTM	major acute phase
12	26	100.0	430	1 KGRTT1	T-kininogen I prec
13	26	100.0	430	2 A23897	major acute phase
14	26	100.0	430	2 B28055	T-kininogen, LMW I
15	26	100.0	431	2 T29186	hypothetical prote
16	26	100.0	433	2 A28055	K-kininogen, LMW I
17	26	100.0	549	2 E83085	conserved hypothet
18	26	100.0	580	2 T21493	hypothetical prote
19	26	100.0	639	2 A25486	kininogen, HMW I p
20	26	100.0	690	2 S64646	biotin-lactetyl-CoA
21	26	100.0	1048	1 XPBEA9	large structural p
22	26	100.0	1068	2 E81965	probable outer mem
23	26	100.0	1082	2 H81020	serotype-1-specific
24	26	100.0	1146	2 B70376	reverse gyrase -A
25	26	100.0	1635	2 T14075	chitinase (EC 3.2.
26	26	100.0	1715	2 JF0128	Bombyx mori recept
27	26	100.0	1885	2 JQ2183	hypothetical 216.5
28	26	100.0	1897	2 C86478	protein F1504.13 [
29	25	96.2	1134	2 F72250	ribosomal protein

30	25	96.2	163	2	AG2365	hypothetical prote
31	25	96.2	197	1	B69131	conserved hypothet
32	25	96.2	214	2	T20688	hypothetical prote
33	25	96.2	248	2	A80177	probable short cha
34	25	96.2	261	2	A83829	thiamin biosynthes
35	25	96.2	268	2	H70118	glucosamine-6-phos
36	25	96.2	319	2	F84966	thioredoxin-disulf
37	25	96.2	326	2	B75101	probable ATP-bindi
38	25	96.2	330	2	T50135	probable l-lactate
39	25	96.2	330	2	E95503	protein F9C16.11 [
40	25	96.2	402	2	B75397	acetyl-CoA acetyl
41	25	96.2	425	1	S48469	probable membrane
42	25	96.2	468	2	E96673	similar to CG1-13
43	25	96.2	482	2	B69803	metabolite transpo
44	25	96.2	577	2	S69631	STPI protein - yea
45	25	96.2	600	2	C86468	probable auxin res
46	25	96.2	615	2	D86473	69.8K hypothetical
47	25	96.2	619	2	C86467	hypothetical prote
48	25	96.2	620	2	E86468	protein F12K21.26
49	25	96.2	637	2	AH1519	hypothetical prote
50	25	96.2	655	2	B83577	transesterolase PA05
51	25	96.2	698	2	A57644	Na+/H+-exchanging
52	25	96.2	754	2	T45910	hypothetical prote
53	25	96.2	767	2	G86476	protein F1504.37 [
54	25	96.2	809	2	A46747	Na+/H+-exchanging
55	25	96.2	813	2	A46748	Na+/H+-exchanging
56	25	96.2	846	2	S59262	hypothetical prote
57	25	96.2	866	2	A11486	phosphoenolpyruvat
58	25	96.2	1038	1	B42641	kinasin-related pr
59	25	96.2	1125	1	F70177	transcription-repa
60	25	96.2	1145	2	T51546	cellulose synthase
61	25	96.2	2037	2	T16881	hypothetical prote
62	25	96.2	3329	2	T42205	breast cancer susc
63	25	96.2	3329	2	T30904	breast cancer tumo
64	25	96.2	3343	2	T42207	breast cancer susc
65	23	88.5	58	2	A45824	hypothetical prote
66	23	88.5	100	2	AB2945	conserved hypothet
67	23	88.5	102	2	F84067	hypothetical prote
68	23	88.5	106	2	G82795	hypothetical prote
69	23	88.5	107	2	H98337	hypothetical prote
70	23	88.5	115	2	F72716	hypothetical prote
71	23	88.5	117	2	T02603	hypothetical prote
72	23	88.5	120	2	A96983	hypothetical prote
73	23	88.5	124	2	AF0946	hypothetical prote
74	23	88.5	135	2	S34546	H+-transporting tw
75	23	88.5	141	2	F64214	histidine triad pr

ALIGNMENTS

RESULT 1

G87164
hypothetical protein [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: G87164
R: Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.
R: Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Hol-
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; PMID:21128732; PMID:11234002
A:Accession: G87164
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-73 <STO>
A:CROSS-references: GB:AL450380; NID:gi13093657; PIDN:CAC30999.1; GSPDB:GN00147
C:Genetics:
A:Gene: ML2044

Query Match 100.0%; Score 26; DB 2; Length 73;

Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
|||||
Db 32 KHAWE 36

RESULT 2
C69193
ribosomal protein L22 - Methanobacterium thermoautotrophicum (strain Delta H)
N/Alternate names: eukaryotic ribosomal protein L17; prokaryotic ribosomal protein L22 H
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 13-Aug-1999
C/Accession: C69193
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: C69193
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-153 <MTH>
A/Cross-references: GB:AE000795; GB:AE000666; NID:G2621036; PIDN:AB84527.1; PID:G262105
A/Experimental source: strain Delta H
C/Genetics:
A/Gene: MTH7
C/Superfamily: rat ribosomal protein L17

Query Match 100.0%; Score 26; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
|||||
Db 29 KHAWE 33

RESULT 3
G69489
LSU ribosomal protein L22P (rp122P) homolog - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 13-Aug-1999
C/Accession: G69489
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.;
Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: G69489
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-155 <KLE>
A/Cross-references: GB:AE000971; GB:AE000782; NID:G2689294; PIDN:AB89352.1; PID:G264864
C/Superfamily: rat ribosomal protein L17

Query Match 100.0%; Score 26; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
|||||
Db 31 KHAWE 35

RESULT 4
H96668
protein FIN19.9 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C/Accession: H96668
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Ki
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: H96668
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-267 <SPO>
A/Cross-references: GB:AE005173; NID:G6633812; PIDN:AAF19671.1; GSPDB:GN00141
C/Genetics:
A/Gene: FIN19.9
A/Map position: 1
C/Superfamily: human 26S proteasome regulatory complex chain p31

Query Match 100.0%; Score 26; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
|||||
Db 143 KHAWE 147

RESULT 5
A82604
oxidoreductase XF2082 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C/Accession: A82604
R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide S
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A83515; MUID:20385717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: A82604
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-268 <STM>
A/Cross-references: GB:AE004023; GB:AE003849; NID:G9107185; PIDN:AAF84881.1; GSPDB:
A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Frasca, S.C.; Franco, M.C.;
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.;
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Marti
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki,
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sa
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Si
M.; Tshukko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Gene: XF2082
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 100.0%; Score 26; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5

Db 149 KHAWE 153
|||||

RESULT 6

T16465
hypothetical protein F55E10.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
C:Accession: T16465
R:Geisels, C.
submitted to the EMBL Data Library, March 1995
A:Description: The sequence of C. elegans cosmid F55E10.
A:Reference number: Z18517
A:Accession: T16465
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-305 <GEI>
A:Cross-references: EMBL:U23455; NID:G746439; PID:G746445; PIDN:AAC46535.1; CESP:F55E10.
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F55E10.6
A:Introns: 5/1; 45/3; 115/3; 141/2; 175/2; 228/1
C:Superfamily: retinol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 100.0%; Score 26; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
|||||
Db 164 KHAWE 168

RESULT 7

A97347
oligopeptide ABC transporter, ATPase component CAC3642 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 19-Jul-2002
C:Accession: A97347
R:Noelling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: A97347
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-340 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK81564.1; PID:G15026742; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3642
C:Superfamily: inner membrane protein malk; ATP-binding cassette homology

Query Match 100.0%; Score 26; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
|||||
Db 130 KHAWE 134

RESULT 8

B82487
hypothetical protein VCA0212 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82487
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Emolaeava, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J. R. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82487

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <HEI>
A:Cross-references: GB:AE004361; GB:AE003853; NID:G9657600; PIDN:AAF96124.1; GSPDB:GN00191
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0212
A:Map position: 2

Query Match 100.0%; Score 26; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 5; Conservative 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
|||||
Db 224 KHAWE 228

RESULT 9

T39939
DNA binding protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39939
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z21892
A:Accession: T39939
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-417 <WOO>
A:Cross-references: EMBL:AL023287; PIDN:CAA18873.1; GSPDB:GN00067; SPDB:SPBC2386.05
A:Experimental source: strain 972h-; cosmid c23E
C:Genetics:
A:Gene: SPDB:SPBC23E.05
A:Map position: 2

Query Match 100.0%; Score 26; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.9e+02; Mismatches 5; Conservative 0; Indels 0; Gaps 0;

Qy 1 KHAWE 5
|||||
Db 223 KHAWE 227

RESULT 10

AC3519
isochoistmate synthase [SC 5.4.99.6] [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Jun-2002
C:Accession: AC3519
R:DeIvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Le Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3519
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-418 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53318.1; PID:G47984204; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10077
A:Map position: 11
C:Superfamily: isochorismate synthase

A23897

major acute phase alpha-1 protein (version 2) - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 20-Aug-1999

C:Accession: A23897; B23897

R:Anderson, K.P.; Heath, E.C.

J. Biol. Chem. 260, 12065-12071, 1985

A:Title: The relationship between rat major acute phase protein and the kininogens.

A:Reference number: A23897; PMID:86008266; PMID:2413019

A:Accession: A23897

A:Molecule type: protein

A:Residues: 1-14 <AND1>

A:Accession: B23897

A:Molecule type: mRNA

A:Residues: 5-430 <AND2>

A:Cross-references: GB:M1661; NID:G205307; PIDN:AAA41570.1; PID:G205308

A:Note: the authors translated the codon CTC for residue 410 as Arg, CTA for residue 415

C:Superfamily: kininogen; cystatin homology

F:19-130/Domain: cystatin homology <CY1>

F:141-252/Domain: cystatin homology <CV2>

F:263-374/Domain: cystatin homology <CV3>

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 430;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 KHAIVE 5

|||||

Db 160 KHAIVE 164

RESULT 14

B28055

T-kininogen, LMW II precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 12-Dec-1997

C:Accession: B28055; E25486; E28526

R:Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.

J. Biol. Chem. 260, 12054-12059, 1985

A:Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and

inhibitor.

A:Reference number: A92496; PMID:86008264; PMID:2413018

A:Accession: B28055

A:Molecule type: mRNA

A:Residues: 1-430 <FUR>

R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.

J. Biol. Chem. 262, 2190-2198, 1987

A:Title: Differing expression patterns and evolution of the rat kininogen gene family.

A:Reference number: A92625; PMID:87137443; PMID:3029068

A:Accession: E25486

A:Molecule type: DNA

A:Residues: 375-430 <KIT>

R:Enyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.

J. Biol. Chem. 263, 973-979, 1988

A:Title: Purification and characterization of rat T-kininogens isolated from plasma of a

A:Reference number: A92729; PMID:88087226; PMID:3121623

A:Accession: B28526

A:Molecule type: protein

A:Residues: 'E', 20-25, 'MD', 28-48, 376-430 <ENJ>

A:Accession: C28526

A:Molecule type: protein

A:Residues: 'E', 20-48, 376-388, 'R', 390-419, 'ER', 422-430 <EN2>

C:Superfamily: kininogen; cystatin homology

C:Keywords: glycoprotein; pyrogutamic acid

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-430/Product: T-kininogen, LMW II #status experimental <MAT>

F:141-252/Domain: cystatin homology <CY1>

F:263-374/Domain: cystatin homology <CV2>

F:263-374/Domain: cystatin homology <CV3>

F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen

F:82-126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/disulfide bonds: #status

RESULT 17

E83085

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 430;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 KHAIVE 5

|||||

Db 160 KHAIVE 164

RESULT 15

T29186

Hypothetical protein C55C3.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T29186

R:Wessene, J.; Stellyes, L.

submitted to the EMBL data library, April 1996

A:Description: The sequence of C. elegans cosmid C55C3.

A:Reference number: Z20585

A:Accession: T29186

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-431 <WOB>

A:Cross-references: EMBL:U53335; PIDN:AAA96169.1; GSPDB:GN000022; CESP:C55C3.4

A:Experimental source: strain Bristol N2; clone C55C3

C:Genetics:

A:Gene: CESP:C55C3.4

A:Map position: 4

A:Introns: 22/3; 57/3; 182/3; 260/3; 355/3; 391/3

C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; T

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 431;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 KHAIVE 5

|||||

Db 106 KHAIVE 110

RESULT 16

A28055

K-kininogen, LMW I precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 15-Nov-1996

C:Accession: A28055

R:Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.

J. Biol. Chem. 260, 12054-12059, 1985

A:Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin

inhibitor.

A:Reference number: A92496; PMID:86008264; PMID:2413018

A:Accession: A28055

A:Molecule type: mRNA

A:Residues: 1-433 <FUR>

C:Superfamily: kininogen; cystatin homology

C:Keywords: alternative splicing

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-433/Product: K-kininogen, LMW I #status predicted <MAT>

F:19-131/Domain: cystatin homology <CY1>

F:142-253/Domain: cystatin homology <CV2>

F:264-375/Domain: cystatin homology <CV3>

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 433;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 KHAIVE 5

|||||

Db 161 KHAIVE 165

conserved hypothetical protein PA4488 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83085
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brady, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83085
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-549 <STO>
A;Cross-references: GB:AE004862; GB:AE004091; NID:99950716; PIDN:AA07876.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4488
C;Superfamily: Escherichia coli hypothetical protein b2226
Query Match 100.0%; Score 26; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAIVE 5
DB 417 KHAIVE 421
RESULT 18
T21493
hypothetical protein F28D1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21493
R;Baynes, C.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19430
A;Accession: T21493
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-580 <WTL>
A;Cross-references: EMBL:Z70684; PIDN:CAA94597.1; GSPDB:GN00022; CBSP:F28D1.1
A;Experimental source: clone F28D1
C;Genetics:
A;Gene: CBSP:F28D1.1
A;Map position: 4
A;Introns: 127/3; 216/2; 477/3
Query Match 100.0%; Score 26; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAIVE 5
DB 552 KHAIVE 556
RESULT 19
A25486
kininogen, HMW I precursor - rat
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C;Accession: A25486
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443; PMID:3029068
A;Accession: A25486
A;Molecule type: mRNA
A;Residues: 1-639 <KIT>
A;Note: the authors translated the codon CAA for residue 347 as Asn

C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F1-19/Domain: signal sequence #status predicted <SIG>
F19-639/Product: kininogen, HMW I #status predicted <MAT>
F19-131/Domain: cystatin homology <CY1>
F142-253/Domain: cystatin homology <CY2>
F1264-375/Domain: cystatin homology <CY3>
Query Match 100.0%; Score 26; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAIVE 5
DB 161 KHAIVE 165
RESULT 20
S64646
biotin-[acetyl-CoA-carboxylase] ligase (EC 6.3.4.15) [validated] - yeast (Saccharom
N;Alternate names: biotin-apoptin ligase; protein D2140; protein YDL141W
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 23-Aug-1996 #text_change 03-Jun-2002
C;Accession: S64646; S67687; S67688
R;Cronan Jr., J.E.; Wallace, J.C.
FEMS Microbiol. Lett. 130, 221-230, 1995
A;Title: The gene encoding the biotin-apoptin ligase of Saccharomyces cerevisiae
A;Reference number: S64646; MUID:95377607; PMID:7649444
A;Accession: S64646
A;Molecule type: DNA
A;Residues: 1-690 <CRO>
A;Cross-references: EMBL:U27182; NID:9886080; PIDN:AAC49057.1; PID:9886081
R;Saluz, H.P.; Woelfl, S.; Hanemann, V.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67677
A;Accession: S67687
A;Molecule type: DNA
A;Residues: 1-690 <SAL>
A;Cross-references: EMBL:Z74189; NID:91431218; PIDN:CAA98714.1; PID:6253238; PID:91
R;Baron, L.; Legros, Y.; Biteau, N.; Monnet, A.; Granotier, C.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67688
A;Accession: S67688
A;Molecule type: DNA
A;Residues: 1-291 <BAR>
A;Cross-references: EMBL:Z74189; GSPDB:GN00004; MIPS:YDL141W
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:BPL1; MIPS:YDL141W
A;Cross-references: SGD:S0002300; MIPS:YDL141W
A;Map position: 4L
C;Function:
A;Description: EC 6.3.4.15 [validated, MUID:95377607]
A;Note: apo-BCCP of E. coli was biotinylated
C;Keywords: ligase; transmembrane protein
F1458-474/Domain: transmembrane #status predicted <TMM>
Query Match 100.0%; Score 26; DB 2; Length 690;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAIVE 5
DB 17 KHAIVE 21
RESULT 21
XP8E9
large structural phosphoprotein pp150 - human cytomegalovirus (strain AD169)
C;Species: human cytomegalovirus, human herpesvirus 5
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 16-Jul-1999
C;Accession: A29533; S09795

A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: H81020
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 11082 <TET>
A;Cross-references: GB|AE002545; GB|AE002098; NID:G7272729; PIDN:AAF42298.1; PID:97
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NME1969

```

Query Match      100.0%; Score 26; DB 2; Length 1082;
Best Local Similarity 100.0%; Pred. NO. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  KXAVE 5
        |||||
Db      713  KXAVE 717

```

reverse gyrase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
C:Accession: B70376
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E
V.

A:Accession: B70376
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1146 <AOF>
A:Cross-references: GB:AE000712; NID:g2983411; PIDN:AC07000.1; PID:g2983417; GB:AE000712
A:Experimental source: strain VF5

T14075
chitinase (EC 3.2.1.14) [yellow fever mosquito
C_Species: Aedes aegypti (yellow fever mosquito)
C_Date: 20-Sep-1999 #sequence_revision 20-Sep-1999
C_Accession: T14075
R_de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.

A:Accession: T14075
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1635
A:Cross-references: EMBL:AF026492; NID:G2564720; PID:G2564721; PIDN:AAB81850.1
C:Genetics:
A:Gene: CHT2
A:Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
Query Match 100.0%; Score 26; DB 2; Length 1635;
Best Local Similarity 100.0%; Pred. No. 7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5

Db 1360 KHAVE 1364

Search completed: June 20, 2003, 20:59:04
Job time : 15.7273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 ; Search time 6.02273 Seconds
(without alignments)
34.433 Million cell updates/sec

Title: US-10-105-008-16

Perfect score: 26

Sequence: 1 KHAVE 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100*

Listing first 75 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	153	1 RL22_METTH	O26115 methanobact
2	26	100.0	155	1 RL22_ARCFU	O28359 archaeoglob
3	26	100.0	430	1 KNT1_RAT	P01048 rattus norv
4	26	100.0	430	1 KNT2_RAT	P08932 rattus norv
5	26	100.0	639	1 KNG_RAT	P08934 rattus norv
6	26	100.0	690	1 BPL1_YEAST	P48445 s biotin--p
7	26	100.0	1048	1 P100_HCMVA	P08318 human cytom
8	26	100.0	1885	1 RRPO_ACLSA	P54891 apple chlor
9	25	96.2	134	1 RS9_THEMA	O9x194 thermotoga
10	25	96.2	197	1 Y245_METTH	O26351 methanobact
11	25	96.2	214	1 RL10_CAEEL	O09533 caenorhabdi
12	25	96.2	254	1 H186_LEPIN	O9s4h7 leptospira
13	25	96.2	259	1 THIG_BACHD	O9kcy6 bacillus ha
14	25	96.2	268	1 NAGB_BORBU	O30564 borrelia bu
15	25	96.2	319	1 TRXB_BUCAL	P57399 buchnera ap
16	25	96.2	326	1 Y800_PPRAB	O9uze8 pyrococcus
17	25	96.2	330	1 LDH_SCHPO	O9p7p7 schizosacch
18	25	96.2	425	1 YIK3_YEAST	P44087 saccharomyc
19	25	96.2	482	1 YFIG_BACSU	P54723 bacillus su
20	25	96.2	577	1 STPL_YEAST	O00947 saccharomyc
21	25	96.2	724	1 EP2K_MOUSE	O08796 mus musculu
22	25	96.2	724	1 EP2K_RAT	P05331 rattus norv
23	25	96.2	809	1 NAH2_RABIT	P50482 oryctolagus
24	25	96.2	812	1 NAH2_HUMAN	O9ubyo homo sapien
25	25	96.2	813	1 NAH2_RAT	P48763 rattus norv
26	25	96.2	846	1 SP98_YEAST	P33540 saccharomyc
27	25	96.2	1038	1 C1N8_YEAST	P27895 saccharomyc
28	25	96.2	1125	1 MFD_BORBU	O51568 borrelia bu
29	25	96.2	3329	1 BRIC_MOUSE	P97929 mus musculu
30	23	88.5	58	1 SINI_BACLI	P22755 bacillus li
31	23	88.5	135	1 ATPE_EUGGR	P31477 euglena gra
32	23	88.5	141	1 YHIT_MYCGL	P47378 mycoplasma
33	23	88.5	154	1 RS5_NICPL	O24111 nicotiana p

34	88.5	172	1 IF3_CAMJE	O9pis2 campylobact
35	88.5	173	1 NADM_METTI	O9uxn8 methanolobu
36	88.5	174	1 PRIC_ECOLI	P23862 escherichia
37	88.5	175	1 Y343_MYCPN	P73143 mycoplasma
38	88.5	192	1 CITX_STRPY	P58160 streptococc
39	88.5	197	1 GRPE_LACSK	O87776 lactobacill
40	88.5	197	1 R85_CICAR	O65731 cicor ariet
41	88.5	199	1 YD63_SCHPO	O10315 schizosacch
42	88.5	219	1 C1DB_HUMAN	O9uhd4 homo sapien
43	88.5	220	1 NHAB_PSECL	P27763 pseudomonas
44	88.5	222	1 Y052_METJA	O60359 methanococc
45	88.5	228	1 UREH_BACSB	O07404 bacillus sp
46	88.5	234	1 URE1_HELHE	P42822 helicobacte
47	88.5	236	1 YF52_PYROH	O59220 pyrococcus
48	88.5	242	1 URIG_PASMU	O9cmi6 paeteurella
49	88.5	265	1 RPIA_ARATH	O9zu38 arabidopsis
50	88.5	271	1 VDLG_HELPF	O9zkw1 helicobacte
51	88.5	280	1 YUSZ_BACSU	P37959 bacillus su
52	88.5	284	1 VDLG_HELPF	O05730 helicobacte
53	88.5	292	1 NLA_DROME	O9x2l8 drosophila
54	88.5	314	1 ANX4_FRAAN	P51074 fragaria an
55	88.5	324	1 FABH_RHOCA	P30790 rhodobacter
56	88.5	326	1 TMOE_PSEME	O00460 pseudomonas
57	88.5	338	1 CTE2_HUMAN	O00154 homo sapien
58	88.5	338	1 CTE2_RAT	O64559 rattus norv
59	88.5	341	1 Y665_METJA	O58079 methanococc
60	88.5	346	1 QUEA_NEIME	O91xw4 neisseria m
61	88.5	357	1 IF35_HUMAN	O00303 homo sapien
62	88.5	382	1 ISDF_CAUCR	O9a715 c ispd/ispf
63	88.5	391	1 DXR_RHIME	O92lp6 rhizobium m
64	88.5	406	1 B9P1_CANAL	P43084 candida alb
65	88.5	422	1 PUR2_BACSU	P12039 bacillus su
66	88.5	425	1 KOTA_ECOLI	P22282 escherichia
67	88.5	427	1 B9P2_YEAST	P36049 saccharomyc
68	88.5	434	1 UDG_RICPR	O05973 rickettsia
69	88.5	438	1 ARP6_YEAST	O12509 saccharomyc
70	88.5	471	1 GLTD_ECOLI	P09832 escherichia
71	88.5	478	1 SEP4_HUMAN	O43236 homo sapien
72	88.5	478	1 SEP4_MOUSE	P28661 mus musculu
73	88.5	483	1 NASF_BACSU	P42437 bacillus su
74	88.5	484	1 PAP2_XENLA	P51005 xenopus lae
75	88.5	505	1 AGAR_ALTAT	P13734 alteromonas

ALIGNMENTS

RESULT 1
RL22_METTH STANDARD; PRT; 153 AA.
ID RL22_METTH
AC O26115;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L22P.
GN RPL22P OR MTH7.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RC MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Blakely D., Cook R., Gilbert K.,
RA Aldredge T., Bashirzadeh R., Lum W., Pothier B., Qiu D.,
RA Harrison D., Hoang L., Keagle P., Wang Y., Wierzbowski J., Gibson R.,
RA Spadafora R., Vicare R., Wang Y., Bush D., Safer H., Pietsch S., Church G.M.,
RA McDougall S., Shimer G., Goyal A., Pietsch S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";

RL J. Bacteriol. 179:7135-7155 (1997).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000795; AAB84527.1; --
DR HSPSP; P48286; 1BXE.
DR InterPro; IPR001063; Ribosomal_L22.
DR Pfam; PF00237; Ribosomal_L22; 1.
DR ProDom; PD001032; Ribosomal_L22; 1.
DR TIGRFAMs; TIGR01038; L22 arch; 1.
DR PROSITE; PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 153 AA; 17346 MW; F56FFD3FAD7516A5 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 153;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
DB 29 KHAWE 33

RESULT 2
RL22_ARCFU STANDARD; PRT; 155 AA.
AC Q28359;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L22P.
GN RPL22P OR AFI920.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-P., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Karlavage A.R., Graham D.E., Kyprides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370 (1997).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000971; AAB89352.1; --
DR HSPSP; P48286; 1BXE.

DR TIGR; AFI920; --
DR InterPro; IPR001063; Ribosomal_L22.
DR Pfam; PF00237; Ribosomal_L22; 1.
DR ProDom; PD001032; Ribosomal_L22; 1.
DR TIGRFAMs; TIGR01038; L22 arch; 1.
DR PROSITE; PS00464; RIBOSOMAL_L22; FALSE NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 155 AA; 18026 MW; 644039F71D7E70F7 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 155;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
DB 31 KHAWE 35

RESULT 3
KNT1_RAT STANDARD; PRT; 430 AA.
ID KNT1_RAT
AC P01048; P04081;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE T-kininogen I precursor (Major acute phase protein) (Alpha-1-MAP)
DE (Thiostatin) [Contains: T-kinin].
GN MAP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86008264; PubMed=2413018;
RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT "Primary structures of the mRNAs encoding the rat precursors for
RT bradykinin and T-kinin. Structural relationship of kininogens with
RT major acute phase protein and alpha 1-cysteine proteinase
RT inhibitor.";
RL J. Biol. Chem. 260:12054-12059 (1985).
RN [2]
RP SEQUENCE OF 5-430 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=86008266; PubMed=2413019;
RA Anderson K.P., Heath E.C.;
RT "The relationship between rat major acute phase protein and the
RT kininogens.";
RL J. Biol. Chem. 260:12065-12071 (1985).
RN [3]
RP SEQUENCE OF 7-430 FROM N.A.
RX MEDLINE=85127561; PubMed=2578992;
RA Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT "Major acute phase alpha 1-protein of the rat is homologous to bovine
RT kininogen and contains the sequence for bradykinin: its synthesis is
RT regulated at the mRNA level.";
RL FEBS Lett. 182:57-61 (1985).
RN [4]
RP SEQUENCE OF 1-65 FROM N.A.
RX MEDLINE=87250580; PubMed=2439509;
RA Fung W.-P., Schreiber G.;
RT "Structure and expression of the genes for major acute phase alpha 1-
RT protein (thiostatin) and kininogen in the rat.";
RL J. Biol. Chem. 262:9298-9308 (1987).
CC -1- FUNCTION: KININOGENS ARE PLASMA GLYCOPROTEINS WITH A NUMBER OF
CC FUNCTIONS: (1) AS PRECURSOR OF THE ACTIVE PEPTIDE BRADYKININ THEY
CC EFFECT SMOOTH MUSCLE CONTRACTION, INDUCTION OF HYPOTENSION AND
CC INCREASE OF VASCULAR PERMEABILITY. (2) THEY PLAY A ROLE IN BLOOD
CC COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND
CC FACTOR XI NEXT TO FACTOR XII. (3) THEY ARE INHIBITOR OF THIO
CC PROTEASES.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- INDUCTION: IN RESPONSE TO AN INFLAMMATORY STIMULANT. T-KININOGEN

II SYNTHESIS IS INDUCED AND THE PLASMA CONCENTRATION OF
T-KININOGEN I IS RAISED.
-1- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
KALLIKREIN.
-1- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
HMW AND LMW KININOGENS PRODUCED BY ALTERNATIVE SPLICING OF THE
SAME GENE, AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND T-II.
-1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
-1- CAUTION: IN ADDITION TO THE CONFLICTS DESCRIBED IN THE FEATURE
TABLE, REF.2. SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS: 257,
262,268,269,295,314,315,331,332, AND 389. IN ALL THOSE POSITIONS
THE ALTERNATE AMINO-ACID IS THE ONE PRESENT IN T-II KININOGEN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M11883; AAA41489.1; -
DR EMBL; M11661; AAA41570.1; -
DR EMBL; M16454; AAA41568.1; -
DR EMBL; X02299; CAA26162.1; ALT_SEQ.
DR PIR; A01286; KGR71.
DR PIR; A01285; KGR7W.
DR PIR; A23897; A23897.
DR PIR; A27115; A27115.
DR GlycoSuiteDB; P01048; -
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 3.
DR SMART; SM00043; Cy; 3.
DR PROSITE; PS00287; CYSTATIN; 2.
KW Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT SIGNAL 1 18
FT CHAIN 19 430 KININOGEN, T-I.
FT PEPTIDE 19 375 HEAVY CHAIN.
FT CHAIN 376 386
FT CHAIN 387 430 LIGHT CHAIN.
FT DOMAIN 19 135 CYSTATIN-LIKE 1.
FT DOMAIN 136 257 CYSTATIN-LIKE 2.
FT DOMAIN 258 375 CYSTATIN-LIKE 3.
FT DISULFID 28 404 INTERCHAIN (BY SIMILARITY).
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 107 125 BY SIMILARITY.
FT DISULFID 141 147 BY SIMILARITY.
FT DISULFID 205 217 BY SIMILARITY.
FT DISULFID 228 247 BY SIMILARITY.
FT DISULFID 263 266 BY SIMILARITY.
FT DISULFID 327 339 BY SIMILARITY.
FT DISULFID 350 369 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 26 28 LNC -> MDR (IN REF. 2).
FT CONFLICT 55 55 V -> L (IN REF. 2).
FT CONFLICT 61 61 E -> K (IN REF. 1).
FT CONFLICT 83 83 C -> Y (IN REF. 3).
FT CONFLICT 166 166 S -> F (IN REF. 2 AND 3).
FT CONFLICT 179 181 REV -> TKI (IN REF. 2).
FT CONFLICT 193 193 N -> D (IN REF. 2).
FT CONFLICT 212 212 S -> F (IN REF. 2).
FT CONFLICT 214 214 R -> H (IN REF. 3).
FT CONFLICT 229 229 T -> R (IN REF. 2).
FT CONFLICT 233 233 H -> Y (IN REF. 2).
FT CONFLICT 257 257 E -> S (IN REF. 2).
FT CONFLICT 262 262 N -> K (IN REF. 2).
FT CONFLICT 264 264 R -> F (IN REF. 2).

FT CONFLICT 268 269 RE -> KN (IN REF. 2).
FT CONFLICT 295 295 I -> L (IN REF. 2).
FT CONFLICT 314 315 VI -> TK (IN REF. 2).
FT CONFLICT 331 332 SK -> TN (IN REF. 2).
FT CONFLICT 389 389 R -> Q (IN REF. 2).
FT CONFLICT 414 414 R -> G (IN REF. 2 AND 3).
FT CONFLICT 415 415 A -> L (IN REF. 2).
FT CONFLICT 420 421 DH -> ER (IN REF. 3).
FT CONFLICT 430 430 P -> S (IN REF. 1).
SQ SEQUENCE 430 AA; 47715 MW; FAEBB78FAF4723C3 CRC64;
Query Match 100.0%; Score 26; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
DB 160 KHAVE 164

RESULT 4
KNT2_RAT
ID KNT2_RAT STANDARD; PRT; 430 AA.
AC P08932;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE T-kininogen II precursor (Major acute phase protein) (Alpha-1-MAP)
DE (Thiohistatin) [contains: T-kinin].
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86008264; PubMed=2413018;
RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT "Primary structures of the mRNAs encoding the rat precursors for
bradykinin and T-kinin. Structural relationship of kininogens with
major acute phase protein and alpha 1-cysteine proteinase
inhibitor.";
RT J. Biol. Chem. 260:12054-12059 (1985).
CC -1- FUNCTION: KININOGENS ARE PLASMA GLYCOPROTEINS WITH A NUMBER OF
FUNCTIONS: (1) AS PRECURSOR OF THE ACTIVE PEPTIDE BRADYKININ THE
EFFECT SMOOTH MUSCLE CONTRACTION. INDUCTION OF HYPOTENSION AND
INCREASE OF VASCULAR PERMEABILITY. (2) THEY PLAY A ROLE IN BLOOD
COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND
FACTOR XI NEXT TO FACTOR XII. (3) THEY ARE INHIBITOR OF THIO
PROTEASES.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- INDUCTION: IN RESPONSE TO AN INFLAMMATORY STIMULANT. T-KININOGEN
II SYNTHESIS IS INDUCED AND THE PLASMA CONCENTRATION OF
T-KININOGEN I IS RAISED.
CC -1- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
KALLIKREIN.
CC -1- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
HMW AND LMW KININOGENS PRODUCED BY ALTERNATIVE SPLICING OF THE
SAME GENE, AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND T-II.
CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M11885; AAA41491.1; -
DR PIR; B28055; B28055.
DR GlycoSuiteDB; P08932; -

DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 3.
 DR SMART; SM00043; CY; 3.
 DR PROSITE; PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
 KW Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 430 KININOGEN, T-II.
 FT CHAIN 19 375 HEAVY CHAIN.
 FT PEPTIDE 376 386 T-KININ.
 FT CHAIN 387 430 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 FT DOMAIN 258 375 CYSTATIN-LIKE 3.
 FT DISULFID 28 404 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 107 125 BY SIMILARITY.
 FT DISULFID 141 144 BY SIMILARITY.
 FT DISULFID 205 217 BY SIMILARITY.
 FT DISULFID 228 247 BY SIMILARITY.
 FT DISULFID 263 266 BY SIMILARITY.
 FT DISULFID 327 339 BY SIMILARITY.
 FT DISULFID 350 369 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 430 AA; 47524 MW; 43EDF02D1BF55076 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 430;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
 Db 160 KHAVE 164

RESULT 5
 KNG_RAT
 ID KNG_RAT STANDARD; PRT; 639 AA.
 AC P08934; P08933;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kininogen precursor [Contains: Bradykinin].
 GN KNG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RX MEDLINE=8713743; PubMed=2413068;
 RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
 RA "Differing expression patterns and evolution of the rat kininogen
 gene family.";
 RL J. Biol. Chem. 262:2190-2198(1987).
 RN [2]
 RP SEQUENCE FROM N.A. (LMW ISOFORM).
 RX MEDLINE=86008264; PubMed=2413018;
 RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
 RA "Primary structures of the mRNAs encoding the rat precursors for
 bradykinin and T-kinin. Structural relationship of kininogens with
 major acute phase protein and alpha 1-cysteine proteinase
 inhibitor.";
 RL J. Biol. Chem. 260:12054-12059(1985).
 RN [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=Buffalo;
 RX MEDLINE=87250580; PubMed=2439509;
 RA Fung W.-P., Schreiber G.;

RT "Structure and expression of the genes for major acute phase alpha 1-
 RT protein (thioestatin) and kininogen in the rat.";
 RL J. Biol. Chem. 262:9298-9308(1987).
 RN [4]
 RP SEQUENCE OF 1-41 FROM N.A.
 RC STRAIN=Mistar; TISSUE=Liver;
 RX MEDLINE=87137465; PubMed=3818598;
 RA Kageyama R., Kitamura N., Okubo H., Nakanishi S.;
 RA "Differing utilization of homologous transcription initiation sites
 of rat K and T kininogen genes under inflammation condition.";
 RL J. Biol. Chem. 262:2345-2351(1987).
 CC -I- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- TISSUE SPECIFICITY: PLASMA.
 CC -I- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -I- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
 CC HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
 CC T-II.
 CC -I- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L29428; AAA41486.1; -
 CC EMBL; M11884; AAA41487.1; -
 CC EMBL; M14369; AAA41484.1; -
 CC EMBL; M14369; AAA41485.1; ALT_SEQ.
 CC EMBL; M16455; AAA41482.1; -
 CC FIR; A25486; A25486.
 CC FIR; A28055; A28055.
 CC InterPro; IPR000010; Cystatin.
 CC InterPro; IPR002395; Kininogen.
 CC Pfam; PF00031; cystatin; 3.
 CC PRINTS; PR00334; KININOGEN.
 CC SMART; SM00043; CY; 3.
 CC PROSITE; PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing; Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 639 KININOGEN.
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
 FT PEPTIDE 381 389 BRADYKININ.
 FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 FT DOMAIN 439 514 HIS-RICH.
 FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.

FT DISULFID 107 126 BY SIMILARITY.
 FT DISULFID 142 145 BY SIMILARITY.
 FT DISULFID 206 218 BY SIMILARITY.
 FT DISULFID 229 248 BY SIMILARITY.
 FT DISULFID 264 267 BY SIMILARITY.
 FT DISULFID 328 340 BY SIMILARITY.
 FT DISULFID 351 370 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 402 433 VPSVIARVQERDPCNEQPHGHGWLHAKO -> RLINS
 CEYKGLLKAGAGPAPERQAEASTVTP (IN ISOFORM LMW).
 FT VARSPLIC 434 639 MISSING (IN ISOFORM LMW).
 FT CONFLICT 61 61 E -> K (IN REF. 2).
 SQ SEQUENCE 639 AA; 70933 MW; D3172DF94FF56AF5 CRC64;
 Query Match 100.0%; Score 26; DB 1; Length 639;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVE 5
 DB 161 KHAVE 165
 RESULT 6
 BPL1 YEAST STANDARD; PRT; 690 AA.
 AC P48445;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Biotin--protein ligase (EC 6.3.4.-) (Biotin Apo-protein ligase)
 DE [Includes: Biotin--[methylmalonyl-CoA-carboxyltransferase] ligase
 DE (EC 6.3.4.9); Biotin--[propionyl-CoA-carboxylase (ATP-hydrolyzing)]
 DE ligase (EC 6.3.4.10) (Holocarboxylase synthetase) (HCS);
 DE Biotin--[methylcrotonyl-CoA-carboxylase] ligase (EC 6.3.4.11);
 DE Biotin--[acetyl-CoA-carboxylase] ligase (EC 6.3.4.15)].
 GN BPL1 OR ACC2 OR VDL141W OR D2140.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=95377607; PubMed=7649444;
 RA Cronan J.E. Jr., Wallace J.C.;
 RT "The gene encoding the biotin- α -apoprotein ligase of *Saccharomyces cerevisiae*.";
 RT FEMS Microbiol. Lett. 130:221-230 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1679;
 RX MEDLINE=97127826; PubMed=8972577;
 RA Woelfl S., Haneman V., Saluz H.P.;
 RT "Analysis of a 26,756 bp segment from the left arm of yeast chromosome IV.";
 RT Yeast 12:1549-1554 (1996).
 RN [3]
 RP SEQUENCE OF 1-291 FROM N.A.
 RA Baron L., Legros Y., Biteau N., Monnet A., Granotier C.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POSTTRANSLATIONAL MODIFICATION OF SPECIFIC PROTEIN BY
 CC ATTACHMENT OF BIOTIN. ACTS ON VARIOUS CARBOXYLASES SUCH AS ACETYL-
 CC COA-CARBOXYLASE, PYRUVATE CARBOXYLASE, PROPIONYL COA CARBOXYLASE,
 CC AND 3-METHYLCROTONYL COA CARBOXYLASE.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[methylmalonyl-CoA:pyruvate
 CC carboxyltransferase] = AMP + diphosphate + [methylmalonyl-

CC COA:pyruvate carboxyltransferase].
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[propanoyl-CoA:carbon-
 CC dioxide ligase (ADP-forming)] = AMP + diphosphate + [propanoyl-
 CC CoA:carbon-dioxide ligase (ADP-forming)].
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[3-methylcrotonyl-
 CC CoA:carbon-dioxide ligase (ADP-forming)] = AMP + diphosphate + [3-
 CC methylcrotonyl-CoA:carbon-dioxide ligase (ADP-forming)].
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[acetyl-CoA:carbon-dioxide
 CC ligase (ADP-forming)] = AMP + diphosphate + [acetyl-CoA:carbon-
 CC dioxide ligase (ADP-forming)].
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SIMILARITY: WITH E.COLI BIRA AND OTHER EUKARYOTIC BIOTIN--PROTEIN
 CC LIGASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U27182; AAC49057.1; -;
 DR EMBL; X96876; CAA65617.1; -;
 DR EMBL; Z74189; CAA98714.1; -;
 DR SGD; S0002300; BPL1.
 DR InterPro; IPR003142; BPL C.
 DR InterPro; IPR004143; BPL_LipA_LipB.
 DR Pfam; PF02237; BPL_C; 1.
 DR Pfam; PF03099; BPL_LipA_LipB; 1.
 KW Ligase.
 SQ SEQUENCE 690 AA; 76363 MW; ADA6BFE411C656AB CRC64;
 Query Match 100.0%; Score 26; DB 1; Length 690;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KHAVE 5
 DB 17 KHAVE 21
 RESULT 7
 P100 HCMVA STANDARD; PRT; 1048 AA.
 ID AC P08318;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Large structural phosphoprotein (PPI50) (150 kDa matrix
 DE phosphoprotein) (Basic phosphoprotein) (BPP).
 GN UL32.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87198858; PubMed=3033266;
 RA Jahn G., Kouzarides T., Mach M., Scholl B.-C., Plachter B.,
 RA Traupe B., Preddle E., Satchwell S.C., Fleckenstein B., Barrell B.G.;
 RT "Map position and nucleotide sequence of the gene for the large
 RT structural phosphoprotein of human cytomegalovirus.";
 RT J. Virol. 61:1358-1367 (1987).
 RN [2]
 RP COMPLETE GENOME.
 RX MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horenell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddle E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169.";
 RT Curr. Top. Microbiol. Immunol. 154:125-169 (1990).


```

[3]
RN CARBOHYDRATE-LINKAGE SITES.
RP MEDLINE=95056066; PubMed=7966627;
RX Greib K.D., Gibson W.G., Hart G.W.;
RT "Site-specific glycosylation of the human cytomegalovirus tegument
RL basic phosphoprotein (UL32) at serine 921 and serine 952.";
RL J. Virol. 68:8339-8349(1994).
CC -1- PPM: MULTIPLY PHOSPHORYLATED.
CC -1- SIMILARITY: TO THE LARGE STRUCTURAL PHOSPHOPROTEINS OF HSV-6 AND
CC HSV-7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17403; CAA35431.1; -.
DR EMBL; M16022; AAA45992.1; -.
DR PIR; A29533; XPBEA9.
DR PIR; S09795; S09795.
DR GlycoSuiteDB; P08318; -.
KW Matrix protein; Phosphorylation; Glycoprotein.
FT CARBOHYD 921 921 O-LINKED (GLCNAC).
FT CARBOHYD 952 952 /FTID=CAR_000179.
FT CARBOHYD 952 952 /FTID=CAR_000195.
FT SEQUENCE 1048 AA; 112688 MW; 9C4DD50B4F56FF1D CRC64;
Query Match 100.0%; Score 26; DB 1; Length 1048;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAWE 5
DB 255 KHAWE 259
|||||
RESULT 8
RPPO ACLSA STANDARD; PRT; 1885 AA.
AC P54891;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (RNA replicase) (216.5 kDa
DE protein) (ORF1).
OS Apple chlorotic leaf spot virus (isolate apple) (ACLSV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Trichovirus.
OX NCBI_TaxID=73472;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93389448; PubMed=8376968;
RA Sato K., Yoshikawa N., Takahashi T.;
RT "Complete nucleotide sequence of the genome of an apple isolate of
RT apple chlorotic leaf spot virus.";
RL J. Gen. Virol. 74:1927-1931(1993).
CC -1- FUNCTION: PROBABLY INVOLVED IN VIRAL RNA REPLICATION.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
CC [RNA] (N).
CC -1- SIMILARITY: TO THE PUTATIVE REPLICATION-ASSOCIATED PROTEINS OF
CC POTEXVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; D14996; BAA03641.1; -.
DR InterPro; IPR005123; 2OG-Fell Oxy.
DR InterPro; IPR001788; RNA dep RNAPol2.
DR InterPro; IPR000606; Viral_helicase1.
DR Pfam; PF00978; RNA dep RNAPol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
DR Pfam; PF03171; 2OG-Fell Oxy; 1.
KW Transferase; RNA-directed RNA polymerase; Nucleotide-binding;
KW Helicase; ATP-binding; RNA replication; Polyprotein.
FT NP_BIND 1060 1067 ATP (POTENTIAL).
FT DOMAIN 1640 1728 POLYMERASE ACTIVE SITE.
FT SEQUENCE 1885 AA; 216505 MW; 7A3FD14D0DBF90E CRC64;
Query Match 100.0%; Score 26; DB 1; Length 1885;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAWE 5
DB 817 KHAWE 821
|||||
RESULT 9
RS9 THEMA STANDARD; PRT; 134 AA.
AC Q9XIG4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S9.
DE RPSI OR TM1453.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Winn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A5001797; AAD36521.1; -.
DR TIGR; TM1453; -.
DR InterPro; IPR000754; Ribosomal S9.
DR Pfam; PF00380; Ribosomal S9; 1.
DR ProDom; PD001627; Ribosomal_S9; 1.
DR PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW Ribosomal protein; Complete proteome.
FT SEQUENCE 134 AA; 15198 MW; 698FB55B73E0A539 CRC64;
Query Match 96.2%; Score 25; DB 1; Length 134;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAWE 5
|||||

```


DB 50 KHAIE 54

RESULT 10
Y249 METHH STANDARD; PRT; 197 AA.
ID Y249 METHH STANDARD; PRT; 197 AA.
AC Q2635L;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Conserved protein.
DE DE
GN MTH249.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shmer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. YCKF SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL: AE000812; AAB84755.1; --
DR InterPro; IPR001347; SIS.
DR Pfam; PF01380; SIS; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 197 AA; 21570 MW; 9840D27964B6F267 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 197;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAIE 5
|||:
16 KHAIE 20

RESULT 11
RL10 CAEEL STANDARD; PRT; 214 AA.
ID RL10 CAEEL STANDARD; PRT; 214 AA.
AC Q09533;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L10 (QM protein homolog).
DE RPL-10 OR P10B5.1.
GN Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdittida; Rhabdittidae;
OC Rhabdittidae; Pelodinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Simms M.;
RA Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

-1- SIMILARITY: BELONGS TO THE LI0E FAMILY OF RIBOSOMAL PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL: Z48334; CAA88308.1; --
DR WormPep; F1085.1; CE01543.
DR InterPro; IPR001197; Ribosomal_L10E.
DR Pfam; PF00826; Ribosomal_L10E; 1.
DR TIGRFAMs; TIGR00279; L10E; 1.
DR PROSITE; PS01257; RIBOSOMAL_L10E; 1.
KW Ribosomal protein.
SQ SEQUENCE 214 AA; 24749 MW; BC12D2B64BD31AAC CRC64;

Query Match 96.2%; Score 25; DB 1; Length 214;
Best Local Similarity 80.0%; Pred. No. 54;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAIE 5
|||:
146 KHAIE 150

RESULT 12
HIS6 LEFIN STANDARD; PRT; 254 AA.
ID HIS6 LEFIN STANDARD; PRT; 254 AA.
AC Q954H7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Imidazole glycerol phosphate synthase subunit hisF (EC 4.1.3.-) (IGP synthase cyclase subunit) (IGP synthase subunit hisF) (IMGP synthase subunit hisF) (IGPS subunit hisF).
GN HISF.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiaceae; Leptospira.
OX NCBI_TaxID=173;
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=Seroovar Hardjo / isolate Hardjoprajitno;
RX MEDLINE=99403407; PubMed=10474199;
RA De la Pena-Moctezuma A., Bulach D.M., Kalambaheti T., Adler B.;
RT Comparative analysis of the LPS biosynthetic loci of the genetic subtypes of serovar Hardjo: Leptospira interrogans subtype Hardjoprajitno and Leptospira borgpetersenii subtype Hardjovovis.";
RL FEMS Microbiol. Lett. 177:319-326(1999).
CC -1- FUNCTION: IGPS catalyzes the conversion of PRFAr and glutamine to IGP, AICAr and glutamate. The hisF subunit catalyzes the cyclization activity that produces IGP and AICAr from PRFAr using the ammonia provided by the his subunit (By similarity).
CC -1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulo-1-ylamino)methyl]eneamine)-1-(5-phospho-glycerol phosphate + 5-carboxamide + l-glutamine = imidazole-glycerol phosphate + 5-aminoimidazol-4-carboxamide ribonucleotide + l-glutamate + H(2)O.
CC -1- PATHWAY: Histidine biosynthesis; fifth step.
CC -1- SUBUNIT: Heterodimer of hisH and hisF (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HISA / HISF FAMILY.
CC -1- CAUTION: The potential active site Asp residue in position 11 is replaced by a Leu.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```
CC EMBL; AF144879; AAD52155.1; -.
DR InterPro; IPR003009; FMN_enzyme.
DR InterPro; IPR000570; His_biosynth.
DR Pfam; PF00977; His_biosynth; 1.
KW Histidine biosynthesis; Lyase.
FT ACT_SITE 130 130.
SQ SEQUENCE 254 AA; 28086 MW; 5447C803CFDDAP5 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 254;
Best Local Similarity 80.0%; Pred. No. 65;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KHAIVE 5
Db 160 KHAIE 164

RESULT 13
THIG_BACHD STANDARD; PRT; 259 AA.
AC Q9KCY6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thiazole biosynthesis protein thig.
GN THIG OR BH1433.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: Required for the synthesis of the thiazole moiety of
CC thiamine (By similarity).
CC -1- COFACTOR: FMN (Potential).
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE THIG FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001512; BAB05152.1; ALT_INIT.
DR InterPro; IPR003009; FMN_enzyme.
KW Thiamine biosynthesis; Flavoprotein; FMN; Complete proteome.
SQ SEQUENCE 259 AA; 27627 MW; 0C52A019EC3EE2D7 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 259;
Best Local Similarity 80.0%; Pred. No. 66;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KHAIVE 5
Db 226 KHAIE 230

RESULT 14
NAGB_BORBU STANDARD; PRT; 268 AA.
ID NAGB_BORBU
AC O30564;
```

```
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucosamine-6-phosphate isomerase (EC 3.5.99.6) (Glucosamine-6-
DE phosphate deaminase) (GNPDA) (GlcN6P deaminase).
GN NAGB OR BB0152.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=1139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SH-2-82;
RA Whitehouse C.A., Austin F.E.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -1- CATALYTIC ACTIVITY: D-glucosamine 6-phosphate + H(2)O = D-fructose
CC 6-phosphate + NH(3).
CC -1- PATHWAY: N-acetyl glucosamine utilization.
CC -1- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC ISOMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF011226; AAB65253.1; -.
DR EMBL; AF001126; AAC66538.1; -.
DR HSSP; P09375; IDEA.
DR TIGR; BB0152; -.
DR InterPro; IPR000457; Glucosamine_iso.
DR InterPro; IPR004547; NagB.
DR Pfam; PF01182; Glucosamine_iso; 1.
DR TIGRFAMs; TIGR00502; nagB; 1.
DR PROSITE; PS01161; GLC_GALNAC_ISOMERASE; 1.
KW Hydrolase; Carbohydrate metabolism; Complete proteome.
FT ACT_SITE 72 72 BY SIMILARITY.
SQ SEQUENCE 268 AA; 30357 MW; 9807089FC4BD75 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 268;
Best Local Similarity 80.0%; Pred. No. 68;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KHAIVE 5
Db 213 KHAIE 217

RESULT 15
TRXB_BUCAI STANDARD; PRT; 319 AA.
ID TRXB_BUCAI
AC P57399;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thioredoxin reductase (EC 1.6.4.5) (TRXR).
```

```

GN TRXB OR BU314.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum)
OS symbiotic bacterium)
OS Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=1093077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT 'Genome sequence of the endocellular bacterial symbiont of aphids
RL Buchnera sp. APS.';
RL Nature 407:181-86(2000).
CC -1- CATALYTIC ACTIVITY: NADPH + oxidized thioredoxin = NADP(+) +
CC reduced thioredoxin.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-II.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF001119; BAB13022.1; --
DR HSP; P09625; 1CLO.
DR InterPro; IPR00759; Adrndx_reductase.
DR InterPro; IPR001327; FAD_pyr_redux.
DR InterPro; IPR001100; Pyr_redux.
DR InterPro; IPR000103; Pyridine_redux_2.
DR Pfam; PF00070; pyr_redux; 1.
DR PRINTS; PR00419; ADXRDXTASE.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; FNDRTASEII.
DR PRINTS; PR00469; FNDRTASEII.
DR ProDom; PD000139; FAD_pyr_redux; 1.
DR TIGRfam; TIGR01292; Trx_redux; 1.
DR PROSITE; PS00573; PYRIDINE_REDOX_2; 1.
KW Redox-active center; Oxidoreductase; NADP; Flavoprotein; FAD;
KW Complete proteome.
FT NP_BIND 7 22 FAD (ADP PART) (PROBABLE).
FT DISULFID 136 139 REDOX-ACTIVE.
FT NP_BIND 148 162 NAD(P) (BY SIMILARITY).
FT NP_BIND 278 298 FAD (FLAVIN PART) (BY SIMILARITY).
SQ SEQUENCE 319 AA; 35105 MW; 27095618CCF1E0F6 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 319;
Best Local Similarity 80.0%; Pred. No. 82;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
DB 71 KHAIE 75

RESULT 16
Y800_PYRAB
ID Y800_PYRAB STANDARD; PRT; 326 AA.
AC Q9UZE8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ATP-binding protein PAB0800.
GN PAB0800.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.

```

```

OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RL structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: BELONGS TO THE ARCHAEL ATPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AJ248286; CAB50111.1; --
DR InterPro; IPR002576; Archaeal_ATPase.
DR Pfam; PF01637; Archaeal_ATPase; 1.
DR ProDom; PD003808; Archaeal_ATPase; 1.
DR Hypoetical protein; ATP-binding; Complete proteome.
KW NP_BIND 28 35 ATP (POTENTIAL).
SQ SEQUENCE 326 AA; 38463 MW; 5647A48265E85704 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 326;
Best Local Similarity 80.0%; Pred. No. 84;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
DB 312 KHAIE 316

RESULT 17
LDH_SCHPO
ID LDH_SCHPO STANDARD; PRT; 330 AA.
AC Q9P7P7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable L-lactate dehydrogenase (EC 1.1.1.27) (L-LDH).
GN SPAC186.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., William R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell J., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
RA Welljens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl U., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu S., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

```

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipakowski G.V., Ussery D., Barrell B.G., Nurese P.,
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC -1- PATHWAY: Anaerobic glycolysis; final step.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL157811; CAB75872.1; -.
DR HSSP; P00344; 2LDB.
DR InterPro; IPR001557; L LDH.
DR InterPro; IPR00205; NAD binding.
DR InterPro; IPR00594; ThiF_domain.
DR InterPro; IPR001236; ldh.
DR Pfam; PF00056; ldh; 1.
DR Pfam; PF02866; ldh_C; 1.
DR PRINTS; PR00086; LLDHGRNASE.
DR PROSITE; PS00064; L LDH; 1.
KW Oxidoreductase; NAD; Glycolysis.
FT ACT_SITE 192 192 ACCEPTS A PROTON DURING CATALYSIS
FT ACT_SITE 192 192 (BY SIMILARITY).
SQ SEQUENCE 330 AA; 35591 MW; 5F68D53F3EA7CDC0 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 330;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
Db 321 KHAIE 325
|||||
SEQUENCE FROM N.A.

RESULT 18
YIK3 YEAST
ID YIK3 YEAST STANDARD; PRT; 425 AA.
AC P40487;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 48.3 kDa protein in MOBI-SGAL intergenic region.
GN YII103W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=288c / AB972;
RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churches C.M., Connor N., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentsch S., Hamlyn N., Hornslett T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO YEAST DIPHTHERIA TOXIN RESISTANCE PROTEIN 2 (DPH2).
CC TO S.FOMBE SPAC33F4.15C, SPBC3B8.05, TO C.ELEGANS C09G5.2, C14B1.5
CC AND M.JANNASCHII MJ0483.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z38125; CAA86277.1; -.
DR PIR; S48469; S48469.
DR SGD; S0001365; YII103W.
DR InterPro; IPR002728; Diphthamide_syn.
DR Pfam; PF01866; Diphthamide_syn; 1.
DR ProDom; PD004399; Diphthamide_syn; 1.
DR TIGRFAMs; TIGR00322; diphth2_R; 1.
KW Hypothetical protein.
SQ SEQUENCE 425 AA; 48310 MW; 22E93852E03C6722 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 425;
Best Local Similarity 80.0%; Pred. No. 11e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAWE 5
Db 421 KHAIE 425
|||||
SEQUENCE FROM N.A.

RESULT 19
YF1G BACSU
ID YF1G BACSU STANDARD; PRT; 482 AA.
AC P54723;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical metabolite transport protein yf1g.
GN YF1G.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96262713; PubMed=8704981;
RA Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;
RT "Determination of a 12 kb nucleotide sequence around the 76 degrees
RT region of the Bacillus subtilis chromosome.";
RL Microbiology 142:1417-1421(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloupi G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerth I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C.Y., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser E., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halesch J., Harwood C.R., Henaut A.,
RA Halbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogiwara K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche M., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sorokin A., Tacconi A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin J., Takahashi H., Takagi T., Takahashi K., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.P., Zumbstein E., Yoshikawa H., Danchin A., "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*,"
RL Nature 390:249-256(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; D50543; BAA09111.1; -
DR EMBL; Z99108; CAB12655.1; -
DR Subtilist; BG11854; yfiG.
DR InterPro; IPR003663; CHO_transport.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTRANSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 30 50 1 (POTENTIAL).
FT DOMAIN 51 59 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 60 80 2 (POTENTIAL).
FT DOMAIN 81 92 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 93 113 3 (POTENTIAL).
FT DOMAIN 114 120 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 121 141 4 (POTENTIAL).
FT DOMAIN 142 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 176 5 (POTENTIAL).
FT DOMAIN 177 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 6 (POTENTIAL).
FT DOMAIN 206 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 284 7 (POTENTIAL).
FT DOMAIN 285 301 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 302 322 8 (POTENTIAL).
FT DOMAIN 323 331 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 332 352 9 (POTENTIAL).
FT TRANSMEM 353 373 10 (POTENTIAL).
FT DOMAIN 374 / 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 11 (POTENTIAL).
FT DOMAIN 422 423 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 424 444 12 (POTENTIAL).
FT DOMAIN 445 482 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 482 AA; 52756 MW; 90851C4F4C48BE01 CRC64;
Query Match 96.2%; Score 25; DB 1; Length 482;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
DB 241 KHAIE 245
RESULT 20
ID_STP1_YEAST STANDARD; PRT; 577 AA.
AC Q00947;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Zinc finger protein STP1.
GN STP1 OR BAP1 OR SSV2 OR YDR463W OR D8035.7.

Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_taxonomy=4932;
RN [1]
RP SEQUENCE OF 59-577 FROM N.A.
RX MEDLINE=92269836; PubMed=1589961;
RA Wang S.S., Stanford D.R., Silvers C.D., Hopper A.K.;
RT "STP1, a gene involved in pre-tRNA processing, encodes a nuclear protein containing zinc finger motifs.";
RL Mol. Cell. Biol. 12:2633-2643(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Bero A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M., Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A., Oetner P., Oh C., Patel F.X., Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N., Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED DIRECTLY IN PRE-TRNA SPLICING. INVOLVED IN THE REMOVAL OF INTERVENING SEQUENCES FROM SOME, BUT NOT ALL, FAMILIES OF INTERVENING SEQUENCES-CONTAINING PRE-TRNAS. REGULATES THE EXPRESSION OF BAP2.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS DIFFERENT TYPES OF ZINC FINGER MOTIFS. STRONG, TO YEAST YHR006W.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; M88597; AAA35124.1; -
DR EMBL; U33050; AAB64913.1; -
DR PIR; S31244; S31244.
DR TRANSFAC; T02402; -
DR SGD; S0002871; STP1.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zfc_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
KW Nuclear protein; zinc-finger; RNA-binding; Metal-binding;
KW tRNA processing.
FT ZN_FING 218 240 C2H2-TYPE.
FT ZN_FING 246 281 C2H2-TYPE (ATYPICAL).
FT ZN_FING 298 323 C2HC-TYPE.
SQ SEQUENCE 577 AA; 64868 MW; 1E79F1E85BB72631 CRC64;
Query Match 96.2%; Score 25; DB 1; Length 577;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHAVE 5
DB 239 KHAIE 243
RESULT 21
ID_EF2K_MOUSE STANDARD; PRT; 724 AA.
AC O08796;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Elongation factor 2 kinase (EC 2.7.1.-) (eEF-2 kinase) (eEF-2K)
DE (Calcium/calmodulin-dependent eukaryotic elongation factor-2 kinase).
GN EEF2K.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; Tissue=Spleen;
RX MEDLINE=97289688; PubMed=9144159;
RA Ryazanov A.G., Ward M.D., Mendola C.E., Pavur K.S., Dorovkov M.V.,
RA Wiedmann M., Erdjument-Bromage H., Tempst P., Farmer T.G.,
RA Probst C.R., Germino F.J., Hait W.N.,
RA "Identification of a new class of protein kinases represented by
RT eukaryotic elongation factor-2 kinase."
RL Proc. Natl. Acad. Sci. U.S.A. 94:4884-4889 (1997).
CC -1- FUNCTION: PHOSPHORYLATES EUKARYOTIC ELONGATION FACTOR-2. BINDS
CC CALMODULIN (BY SIMILARITY).
CC INTRAMOLECULAR AUTOPHOSPHORYLATION, AND THIS RESULTS IN IT
CC BECOMING PARTIALLY CALCIUM/CALMODULIN-INDEPENDENT.
CC -1- SUBUNIT: MONOMER OR HOMODIMER (POTENTIAL).
CC -1- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED. PARTICULARLY ABUNDANT
CC IN SKELETAL MUSCLE AND HEART.
CC -1- SIMILARITY: BELONGS TO THE MCK / EF-2 PROTEIN KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; U93848; AAB58271.1;
DR MGD; MGI:1195261; Ref2k.
DR InterPro; IPR004166; MCK_EF2_kinase.
DR Pfam; PF02816; MCK_EF2_kinase; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Calcium;
KW Calmodulin-binding; Phosphorylation.
FT NP BIND 295 301 ATP (BY SIMILARITY).
FT DOMAIN 593 609 CALMODULIN-BINDING (POTENTIAL).
FT DOMAIN 609 626 PSEUDOSUBSTRATE/AUTOINHIBITORY DOMAIN
FT (POTENTIAL).
FT NP BIND 295 301 ATP (BY SIMILARITY).
FT SEQUENCE 724 AA; 81738 MW; B892D7D547C8E7AE CRC64;
Query Match 96.2%; Score 25; DB 1; Length 724;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVE 5
Db 85 KHAIE 89
RESULT 22
ID EF2K RAT STANDARD; PRT; 724 AA.
AC P70531; O09089;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Elongation factor 2 kinase (EC 2.7.1.-) (eEF-2 kinase) (eEF-2K)
DE (Calcium/calmodulin-dependent eukaryotic elongation factor-2 kinase).
GN EF2K.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Skeletal muscle;
RX MEDLINE=96291917; PubMed=8663182;
RA Redpath N.T., Price N.T., Proud C.G.;
RA "Cloning and expression of cDNA encoding protein synthesis elongation
RT factor-2 kinase.";

J. Biol. Chem. 271:17547-17554 (1996).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=New England Deaconess Hospital; Tissue=Pheochromocytoma;
RX MEDLINE=97289688; PubMed=9144159;
RA Ryazanov A.G., Ward M.D., Mendola C.E., Pavur K.S., Dorovkov M.V.,
RA Wiedmann M., Erdjument-Bromage H., Tempst P., Farmer T.G.,
RA Probst C.R., Germino F.J., Hait W.N.,
RA "Identification of a new class of protein kinases represented by
RT eukaryotic elongation factor-2 kinase."
RL Proc. Natl. Acad. Sci. U.S.A. 94:4884-4889 (1997).
CC -1- FUNCTION: IN VITRO, IT PHOSPHORYLATES TWO ADJACENT THREONINE
CC RESIDUES, THR-57 AND THR-59, IN THE N-TERMINUS OF EUKARYOTIC
CC ELONGATION FACTOR-2. BINDS CALMODULIN.
CC -1- ENZYME REGULATION: UNDERGOES CALCIUM/CALMODULIN-DEPENDENT
CC INTRAMOLECULAR AUTOPHOSPHORYLATION, AND THIS RESULTS IN IT
CC BECOMING PARTIALLY CALCIUM/CALMODULIN-INDEPENDENT.
CC -1- SUBUNIT: MONOMER OR HOMODIMER (POTENTIAL).
CC -1- TISSUE SPECIFICITY: MOSTLY IN SKELETAL MUSCLE.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE MCK / EF-2 PROTEIN KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; X96426; CAA65286.1;
DR EMBL; U93849; AAB58272.1;
DR InterPro; IPR004166; MCK_EF2_kinase.
DR Pfam; PF02816; MCK_EF2_kinase; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Calcium;
KW Calmodulin-binding; Phosphorylation.
FT DOMAIN 593 609 CALMODULIN-BINDING (POTENTIAL).
FT DOMAIN 609 626 PSEUDOSUBSTRATE/AUTOINHIBITORY DOMAIN
FT (POTENTIAL).
FT NP BIND 295 301 ATP (BY SIMILARITY).
FT SEQUENCE 724 AA; 81489 MW; E4DD832AE9F7165 CRC64;
Query Match 96.2%; Score 25; DB 1; Length 724;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KHAVE 5
Db 85 KHAIE 89
RESULT 23
ID NAH2_RABIT STANDARD; PRT; 809 AA.
AC P50482;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sodium/hydrogen exchanger 2 (Na(+)/H(+) exchanger 2) (NHE-2).
GN SLC9A2 OR NHE2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white;
RX MEDLINE=93280159; PubMed=7685025;
RA Tee C.-M., Levine S.A., Yun C.H., Montrose M.H., Little P.J.,
RA Pouyssegur J., Donowitz M.;
RA "Cloning and expression of a rabbit cDNA encoding a serum-activated
RT ethyloisopropylamide-resistant epithelial Na+/H+ exchanger isoform
RT (NHE-2).";

RL J. Biol. Chem. 269:11917-11924 (1993).

CC - FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED

CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL

CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD

CC SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN

CC COLONIC SODIUM ABSORPTION.

CC - SUBCELLULAR LOCATION: Integral membrane protein.

CC - TISSUE SPECIFICITY: HIGH LEVELS IN INTESTINE AND KIDNEY.

CC - PTM: PHOSPHORYLATED (POSSIBLE).

CC - SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.

CC - CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC

CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L13733; -; NOT ANNOTATED CDS.

DR InterPro; IPR000676; NaH Exchng.

DR InterPro; IPR004709; NaH Exchng3.

DR Pfam; PF00999; Na H Exchanger; 1.

DR PRINTS; PR01084; NAHEXCHNGR.

DR TIGRFAMs; TIGR00840; b_cpai; 1.

KW Transmembrane; Glycoprotein; transport; Antiport; Sodium transport;

KW Multigene family; Phosphorylation.

FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 17 37 A (M1) (POTENTIAL).

FT DOMAIN 38 79 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 80 100 B (M2) (POTENTIAL).

FT DOMAIN 101 106 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 107 127 C (M3) (POTENTIAL).

FT DOMAIN 128 138 D (M4) (POTENTIAL).

FT TRANSMEM 139 159 E (M5) (POTENTIAL).

FT DOMAIN 160 168 F (M6) (POTENTIAL).

FT TRANSMEM 169 189 G (M7) (POTENTIAL).

FT DOMAIN 190 208 H (M8) (POTENTIAL).

FT TRANSMEM 209 229 I (M9) (POTENTIAL).

FT DOMAIN 230 236 J (M10) (POTENTIAL).

FT TRANSMEM 237 257 K (M11) (POTENTIAL).

FT DOMAIN 258 277 L (M12) (POTENTIAL).

FT TRANSMEM 278 298 M (M13) (POTENTIAL).

FT DOMAIN 299 308 N (M14) (POTENTIAL).

FT TRANSMEM 309 329 O (M15) (POTENTIAL).

FT CARBOHYD 330 350 N-LINKED (GLCNAC...), (POTENTIAL).

SQ SEQUENCE 809 AA; 90744 MW; DBD00B45443DB7A6 CRC64;

Query Match 96.2%; Score 25; DB 1; Length 809;

Best Local Similarity 80.0%; Pred. No. 2.1e-02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KHAVE 5

Db 562 KHAIE 566

RESULT 24

NAH2 HUMAN STANDARD; PRT; 812 AA.

ID NAH2 HUMAN

AC Q9UB70;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DR Sodium/hydrogen exchanger 2 (Na(+)/H(+) exchanger 2) (NHE-2).

GN SLC9A2 OR NHE2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RN SEQUENCE FROM N.A.

RP TISSUE=Colon;

RC MEDLINE=99375108; PubMed=1044453;

RX Malakooti J., Dahdal R.Y., Schmidt L., Layden T.J., Dudeja P.K.,

RA Ramaswamy K.;

RA "Molecular cloning, tissue distribution, and functional expression of

RT the human Na(+)/H(+) exchanger NHE2.";

RL Am. J. Physiol. 277:G383-G390 (1999).

RN [2]

RP SEQUENCE FROM N.A.

RL Hou S., Wohldmann P.;

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

CC - FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED

CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL

CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD

CC SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN

CC COLONIC SODIUM ABSORPTION.

CC - SUBCELLULAR LOCATION: Integral membrane protein.

CC - TISSUE SPECIFICITY: EXPRESSED IN SKELETAL MUSCLE, COLON AND

CC KIDNEY. LOWER LEVELS IN THE TESTIS, PROSTATE, OVARY, AND SMALL

CC INTESTINE.

CC - PTM: PHOSPHORYLATED (POSSIBLE).

CC - SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.

CC - CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC

CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF073299; AADA1635.1; -;

DR EMBL; AC007239; AAF19248.1; -;

DR Genew; HGNC:11072; SLC9A2.

DR MIN; 600530; -;

DR InterPro; IPR000676; NaH Exchng.

DR InterPro; IPR004709; NaH Exchng3.

DR Pfam; PF00999; Na H Exchanger; 1.

DR PRINTS; PR01084; NAHEXCHNGR.

DR TIGRFAMs; TIGR00840; b_cpai; 1.

KW Transmembrane; Glycoprotein; transport; Antiport; Sodium transport;

KW Multigene family; Phosphorylation.

FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 14 33 A (M1) HYDROPHOBIC.

FT DOMAIN 34 179 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 80 100 B (M2) HYDROPHOBIC.

FT DOMAIN 101 106 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 107 127 C (M3) (POTENTIAL).

FT DOMAIN 128 138 D (M4) (POTENTIAL).

FT TRANSMEM 139 159 E (M5) (POTENTIAL).

FT DOMAIN 160 168 F (M6) (POTENTIAL).

FT TRANSMEM 169 189 G (M7) (POTENTIAL).

FT DOMAIN 190 208 H (M8) (POTENTIAL).

FT TRANSMEM 209 229 I (M9) (POTENTIAL).

FT DOMAIN 230 236 J (M10) (POTENTIAL).

FT TRANSMEM 237 257 K (M11) (POTENTIAL).

FT DOMAIN 258 277 L (M12) (POTENTIAL).

FT TRANSMEM 278 298 M (M13) (POTENTIAL).

FT DOMAIN 299 308 N (M14) (POTENTIAL).

FT TRANSMEM 309 329 O (M15) (POTENTIAL).

FT DOMAIN 329 360 EXTRACELLULAR (POTENTIAL).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw.model

Run on: June 20, 2003, 20:42:49 ; Search time 26.1636 Seconds
(without alignments)
39.078 Million cell updates/sec

Title: US-10-105-008-16
Perfect score: 26
Sequence: 1 KHAVE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SPTREMBL_21:

1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriaph:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	73	16 Q9CBF4	Q9cbf4 mycobacteri
2	26	100.0	113	5 O17418	O17418 anopheles s
3	26	100.0	159	10 Q9PHY0	Q9phy0 arabidopsis
4	26	100.0	195	5 Q95051	Q95051 tritrichomo
5	26	100.0	217	16 Q8Y2Z5	Q8y2z5 ralstonia s
6	26	100.0	235	13 Q8Q615	Q8q615 fundulus he
7	26	100.0	248	16 Q987H1	Q987h1 rhizobium 1
8	26	100.0	249	17 Q8T2H2	Q8t2h2 pyrococcus
9	26	100.0	258	2 Q8RPR9	Q8rpr9 bifidobacte
10	26	100.0	262	17 Q8RTP6	Q8rtp6 methanosarc
11	26	100.0	267	10 Q9SCW3	Q9scw3 arabidopsis
12	26	100.0	267	10 Q8W423	Q8w423 oryza sativ
13	26	100.0	268	16 Q9PBQ8	Q9pbq8 xylella fas
14	26	100.0	268	16 Q8XUH9	Q8xuh9 ralstonia s
15	26	100.0	271	16 Q99PY9	Q99py9 streptomyce
16	26	100.0	276	2 Q9ALU7	Q9alu7 enterobacte

17	26	100.0	295	12 Q9PY98	Q9py98 xestia c-ni
18	26	100.0	305	5 Q20840	Q20840 caenorhabdi
19	26	100.0	307	2 Q84916	Q84916 thioabacillu
20	26	100.0	315	17 Q8ZX44	Q8zx44 pyrobaculum
21	26	100.0	326	17 Q8TNZ4	Q8tnz4 methanosarc
22	26	100.0	332	10 Q9FRD8	Q9frd8 oryza sativ
23	26	100.0	340	16 Q97D39	Q97d39 clostridium
24	26	100.0	348	2 Q8VNZ6	Q8vnz6 bordetella
25	26	100.0	349	16 Q9KMW6	Q9kmw6 vibrio chol
26	26	100.0	366	2 Q9F0C2	Q9f0c2 lactobacill
27	26	100.0	392	17 Q8TME0	Q8tme0 methanosarc
28	26	100.0	415	5 Q44228	Q44228 halocynthia
29	26	100.0	417	3 Q60180	Q60180 schizosacch
30	26	100.0	418	16 Q9YDU7	Q9ydu7 bruceella me
31	26	100.0	423	11 P70517	P70517 rattus norv
32	26	100.0	427	2 Q9E280	Q9e280 bruceella ab
33	26	100.0	430	11 Q63581	Q63581 rattus norv
34	26	100.0	431	5 Q18865	Q18865 caenorhabdi
35	26	100.0	432	4 Q9Y2C7	Q9y2c7 homo sapien
36	26	100.0	482	10 Q9LEN5	Q9len5 cicier ariet
37	26	100.0	492	5 Q8T093	Q8t093 drosophila
38	26	100.0	549	16 Q9HVT3	Q9hvt3 pseudomonas
39	26	100.0	580	5 Q19873	Q19873 caenorhabdi
40	26	100.0	590	2 Q93E16	Q93e16 rhizobium 1
41	26	100.0	793	5 Q8WQ84	Q8wq84 drosophila
42	26	100.0	924	5 Q18606	Q18606 caenorhabdi
43	26	100.0	955	5 Q961Q5	Q961q5 drosophila
44	26	100.0	997	5 Q9GT27	Q9gt27 drosophila
45	26	100.0	1039	5 Q9VMG8	Q9vmg8 drosophila
46	26	100.0	1046	12 Q9DXH5	Q9dxh5 human cytom
47	26	100.0	1067	2 Q9AE78	Q9ae78 neisseria m
48	26	100.0	1068	16 Q9JW98	Q9jw98 neisseria m
49	26	100.0	1082	16 Q9JXM7	Q9jxm7 neisseria m
50	26	100.0	1146	16 Q67037	Q67037 aquifex aeo
51	26	100.0	1443	5 Q9VPR0	Q9vpr0 drosophila
52	26	100.0	1592	5 Q9GU17	Q9gu17 lucilia cup
53	26	100.0	1635	5 Q17412	Q17412 aedes aegyp
54	26	100.0	1666	5 Q917P4	Q917p4 drosophila
55	26	100.0	1715	5 Q9XY09	Q9xy09 bombyx mori
56	26	100.0	1716	5 Q9NDN5	Q9ndn5 bombyx mori
57	26	100.0	1716	5 Q9NDN4	Q9ndn4 bombyx mori
58	26	100.0	1716	5 Q9NDN3	Q9ndn3 bombyx mori
59	26	100.0	1723	5 Q95WK9	Q95wk9 lymantria d
60	26	100.0	1732	5 Q963J6	Q963j6 heliothis v
61	26	100.0	1887	10 Q9LQH2	Q9lqh2 arabidopsis
62	26	100.0	2186	5 Q9N906	Q9n906 trypanosoma
63	26	96.2	105	11 Q89399	Q89399 meriones un
64	26	96.2	138	11 Q35224	Q35224 cynomys lud
65	26	96.2	146	12 Q9YUHI	Q9yuh1 rotavirus b
66	26	96.2	163	16 Q8YNT2	Q8ynt2 anabaena sp
67	26	96.2	172	11 Q9WUJ4	Q9wuj4 mus musculu
68	26	96.2	198	4 Q00618	Q00618 homo sapien
69	26	96.2	225	17 Q8TQ54	Q8tq54 methanosarc
70	26	96.2	244	16 Q8RAY6	Q8ray6 thermocoe
71	26	96.2	248	16 Q8ZG63	Q8zge3 yersinia pe
72	26	96.2	257	16 Q8XQV9	Q8xqv9 ralstonia s
73	26	96.2	259	10 Q9M476	Q9m476 dendrobium
74	26	96.2	262	2 Q32601	Q32601 pantoea agg
75	26	96.2	269	17 Q8TR63	Q8tr63 methanosarc

ALIGNMENTS

RESULT 1
Q9CBF4 PRELIMINARY; PRT; 73 AA.
ID Q9CBF4
AC Q9CBF4; (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE Hypothetical protein ML2044.
GN ML2044.

```

Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHAWE 5
          |||||
Db      95 KHAWE 99

RESULT 3
Q9FHY0
ID Q9FHY0 PRELIMINARY; PRT; 159 AA.
AC Q9FHY0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Similarity to 26S proteasome regulatory complex subunit.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

```

RC STRAIN=COLUMBIA;
RX MEDLINE=99397451; PubMed=10470850;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA Miyajima N., Tabata S.,
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT Sequence features of the regions of 1,011,550 bp covered by seventeen
RT p1 and TAC clones.";
RL DNA Res. 6:183-195(1999).
RW EMBL; AB017067; BAB08437.1; -.
KW Proteasome.

```

KW Proteasome.
SQ SEQUENCE 159 AA; 18198 MW; 0594E685BF060673 CRC64;

Query Match 100.0%; Score 26; DB 10; Length 159;
Best Local Similarity 100.0%; Pred. NO. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAVE 5
   |||||
DB 35 KHAVE 39

RESULT 4
Q95051 PRELIMINARY; PRT; 195 AA.
ID Q95051
AC Q95051;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Superoxide dismutase 1 (Iron superoxide dismutase) (EC 1.15.1.1).
OS SOD1.
GN Trichomonas foetus (Trichomonas foetus).
OC Eukaryota; Parabasalides; Trichomonadida; Trichomonadidae;
OC Trichomonas.
NCBI_TaxID=5724;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MTB5-330.1;
RA Granger B.L., Warwood S.J., Hayashi H., Ohashi M.;
RRL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 32-156 FROM N.A.
RC STRAIN=ATCC 30924;
RA Touzel J.;
RRL Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 32-156 FROM N.A.
RC STRAIN=ATCC 30924;
RER MEDLINE=97047378; PubMed=8892298;
RA Viscogliosi E., Durieux I., Delgado-Viscogliosi P., Bayle D., Dive D.;
RT "Phylogenetic implication of iron-containing superoxide dismutase

```

genes from trichomonad species.";

RT Mol. Biochem. Parasitol. 80:209-214 (1996).

CL -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

CC -1- COPACATOR: COPPER AND ZINC OR IRON OR MANGANESE.

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

CC FAMILY.

DR EMBL; U66073; AAC47734.1; -;

DR EMBL; Z70668; CAA94525.1; -;

DR HSSP; P09157; IISC.

DR InterPro; IPR001189; SODismutase.

DR Pfam; PF00081; sodfe; 1.

DR Pfam; PF02777; sodfe; 1.

DR ProDom; PD000475; SODismutase; 1.

KW Oxidoreductase.

SQ SEQUENCE 195 AA; 22171 MW; 002CDD929B6C812B CRC64;

Query Match 100.0%; Score 26; DB 5; Length 195;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
| | | | |

Db 20 KHAVE 24

RESULT 5 PRELIMINARY; PRT; 217 AA.

Q8Y2Z5 Q8Y2Z5

AC Q8Y2Z5, 2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypothetical protein RSC0186.

GN RSC0186 OR RS00617.

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI_TaxID=305;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GMI1000;

RX MEDLINE=21631879; PubMed=11823852;

RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,

RA Arlat M., Ballaut A., Brottier P., Camus J.C., Cattolico L.,

RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,

RA Gaepin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,

RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,

RA Weissbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";

RL Nature 415:497-502(2002)

RR EMBL; AL646057; CAU13714.1; -;

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 217 AA; 23425 MW; 881A0778D5DB419 CRC64;

Query Match 100.0%; Score 26; DB 16; Length 217;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
| | | | |

Db 47 KHAVE 51

RESULT 6 PRELIMINARY; PRT; 235 AA.

Q8QG15 Q8QG15

AC Q8QG15, 2002 (TrEMBLrel. 21, Created)

DT 01-MAR-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 21, Last annotation update)

DE Hypothetical protein alpha-like protein (fragment).

```

Db      157 KHAVE 161
RESULT 8
Q8T2H2
ID      Q8T2H2      PRELIMINARY;      PRT;      249 AA.
AC      Q8T2H2;
DT      01-JUN-2002 (TReMBLrel. 21, Created)
DT      01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT      01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE      Hypothetical protein PF2022.
GN      Pyrococcus furiosus.
OS      Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcaceae;
OC      Pyrococcus
OX      NCBI_TaxID=2261;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=VC1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA      Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT      "The complete sequence of the Pyrococcus furiosus genome.";
RL      Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AE010294; AAU82146.1; -
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 249 AA; 27594 MW; BBBC2F00853C33BF CRC64;

Query Match      100.0%; Score 26; DB 17; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHAVE 5
      |||||
Db      175 KHAVE 179

RESULT 9
Q8RPR9
ID      Q8RPR9      PRELIMINARY;      PRT;      258 AA.
AC      Q8RPR9;
DT      01-JUN-2002 (TReMBLrel. 21, Created)
DT      01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT      01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE      Putative plasmid transfer protein.
GN      TRA.
OS      Bifidobacterium pseudocatenulatum.
OG      Plasmid p4M.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OX      Bifidobacteriales; Bifidobacteriaceae; Bifidobacterium.
OX      NCBI_TaxID=28026;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=VMB4M;
RA      Smeianov V.V., Efimov B.A., Korschunov V.M., Steele J.L.;
RT      "Characterization of plasmid p4M from Bifidobacterium
      pseudocatenulatum.";
RL      Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF359574; AA00236.1; -
KW      Plasmid.
SQ      SEQUENCE 258 AA; 28695 MW; 5D66CA865F646457 CRC64;

Query Match      100.0%; Score 26; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHAVE 5
      |||||
Db      207 KHAVE 211

RESULT 10
Q8TTP6
ID      Q8TTP6      PRELIMINARY;      PRT;      262 AA.
AC      Q8TTP6;

```

```

DT      01-JUN-2002 (TReMBLrel. 21, Created)
DT      01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT      01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE      Hypothetical protein MA0380.
GN      MA0380.
OS      Methanoscarchina acetivorans.
OC      Archaea; Euryarchaeota; Methanococci; Methanoscarchinales;
OX      Methanoscarchinaceae; Methanoscarchina.
OX      NCBI_TaxID=2214;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C2A / ATCC 35395 / DSM 2834;
RX      MEDLINE=21929760; PubMed=11932238;
RA      Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
      Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
      Allen N., Maylor J., Stange-Thomann N., Dearellano K., Johnson R.,
      Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
      Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
      Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
      Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
      Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
      Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
      Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
      Metcalf W.W., Birren B.;
RA      "The genome of Methanoscarchina acetivorans reveals extensive metabolic
      RT      and physiological diversity.";
RL      Genome Res. 12:532-542(2002).
DR      EMBL; AB010697; AA003832.1; -
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 262 AA; 30203 MW; 6F597FEE1E3303FF CRC64;

Query Match      100.0%; Score 26; DB 17; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHAVE 5
      |||||
Db      95 KHAVE 99

RESULT 11
Q9SGW3
ID      Q9SGW3      PRELIMINARY;      PRT;      267 AA.
AC      Q9SGW3;
DT      01-MAY-2000 (TReMBLrel. 13, Created)
DT      01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT      01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE      FIN19.9 (AT1G64520/FIN19.10).
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX      eurobids 11; Brassicales; Brassicaceae; Arabidopsi.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
      Shinn P., Tambunga G., Altafi H., Bei Q., Chin C., Chlou J., Choi E.,
      Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
      Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharezy N.,
      Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
      Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A.,
      Theologis A., Ecker J.R.;
RA      "Genomic sequence for Arabidopsis thaliana BAC F1N19 from chromosome
      RT      I.";
RL      Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
      Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
      Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
      Kawai J., Lam B., Lee J.M., Lin J., Liu S.K., Miranda M., Narusaka M.,
      Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
      Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,

```


RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Axlal M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choigne N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*,"
RL Nature 415:497-502(2002).
DR EMBL; AL646074; CAD16997.1; --
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 268 AA; 28554 MW; F58670DAEE2AA9DE CRC64;

Query Match 100.0%; Score 26; DB 16; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAIVE 5
DB 149 KHAIVE 153

RESULT 15
Q99PY9 PRELIMINARY; PRT; 271 AA.
AC Q99PY9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative short-chain oxidoreductase.
GN SCPI.311C AND SCPI.43.
OS Streptomyces coelicolor.
OG Plasmid SCPI.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Murphy L.D., Harris D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Ikeda K., Yamasaki M., Kinaishi H.;
RT "Cloning and physical mapping of the EcoRI fragments of the giant
linear plasmid SCPI.,"
RL J. Bacteriol. 180:2796-2799(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete *Streptomyces
coelicolor* A3(2).";
RL Nature 417:141-147(2002).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
(SDR) FAMILY.

DR EMBL; AL590464; CAC36836.1; --
DR EMBL; AL590463; CAC36565.1; --
DR HSSP; P14061; 1FDS.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR POSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Oxidoreductase; Plasmid.
SQ SEQUENCE 271 AA; 28446 MW; F1347AB523E495CC CRC64;

Query Match 100.0%; Score 26; DB 16; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHAIVE 5
DB 150 KHAIVE 154

RESULT 16
Q9ALU7 PRELIMINARY; PRT; 276 AA.
AC Q9ALU7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative oxidoreductase.
OS Enterobacter cloacae.
OG Plasmid pPDL12.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=550;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99;
RA MEDLINE=91123219; PubMed=1991731;
RA Eaton R.W., Karns J.S.;
RT "Cloning and comparison of the DNA encoding ammelide aminohydrolase
and cyanuric acid amidohydrolase from three s-triazine-degrading
bacterial strains,"
RL J. Bacteriol. 173:1363-1366(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=99;
RA Karns J.S., Eaton R.W.;
RT "Genes encoding s-triazine degradation are plasmid-borne in *Klebsiella
pneumoniae* strain 99,"
RL J. Agric. Food Chem. 45:1017-1022(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=99;
RA Eaton R.W.;
RT "DNA sequences of s-triazine catabolic gene regions from *Enterobacter
cloacae* strain 99, *Acidovorax avenae* NRRLB-12227, and *Pseudomonas
huttensis* NRRLB-12228,"
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
(SDR) FAMILY.
CC EMBL; AF342826; AAK11695.1; --
DR HSSP; P14061; 1FDS.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002029; Peptidase_S8.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR POSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
KW Oxidoreductase; Plasmid.
SQ SEQUENCE 276 AA; 29762 MW; E13B18C52D6F5F9C CRC64;

Query Match 100.0%; Score 26; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5
 Db 150 KHAVE 154

RESULT 17
 Q9PY8 PRELIMINARY; PRT; 295 AA.
 AC Q9PY8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ORF55.
 OS Xestia c-nigrum granulosis virus (XngV) (Xestia c-nigrum
 granulovirus).
 OC Viruses; daDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
 OX NCBI_TaxID=51677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99434230; PubMed=10502508;
 RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
 RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
 RL Virology 282:277-297(1999).
 DR EMBL; AF162221; AAF05169.1; --
 SQ SEQUENCE 295 AA; 33625 MW; 48D1642F30F97A96 CRC64;

Query Match 100.0%; Score 26; DB 12; Length 295;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5
 Db 182 KHAVE 186

RESULT 18
 Q20840 PRELIMINARY; PRT; 305 AA.
 AC Q20840;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE D-beta-hydroxybutyrate dehydrogenase homolog.
 GN F55E10.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BRISTOL N2;
 RA MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."; Nature 368:32-38(1994).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BRISTOL N2;
 RA Geisel C.;
 DT submitted (APR-1995) to the EMBL/GenBank/DBJ databases.

FROM N.A.
 NCBI_TaxID=12;
 NCBI_TaxID=12;

RA Waterston R.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: SOME TO RETINOL DEHYDROGENASE.
 DR EMBL; U23455; AAC46555.1; --
 DR HSP; P14061; IPDV.
 DR WormPep; F55E10.6; CE02786.
 DR InterPro; IPR002198; ADH short.
 DR Pfam; PF00106; adh short; 1.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 KW Oxidoreductase; NAD; 47 NAD (BY SIMILARITY).
 FT NP_BIND 23 47
 FT ACT_SITE 160 160 BY SIMILARITY.
 SQ SEQUENCE 305 AA; 34456 MW; 9BD7748A66DB0E04 CRC64;

Query Match 100.0%; Score 26; DB 5; Length 305;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5
 Db 164 KHAVE 168

RESULT 19
 O84916 PRELIMINARY; PRT; 307 AA.
 ID O84916
 AC O84916;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 33.7 kDa protein.
 GN CBR.
 OS Thiobacillus intermedius.
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Thiomonas.
 OX NCBI_TaxID=926;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Shively J.M., Soyer F.;
 RT "A Form II Rubisco Gene and Associated Genes in Thiobacillus
 RT intermedius";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.
 DR EMBL; AF012127; AAC24963.1; --
 DR InterPro; IPR000847; HTH_LysR.
 DR InterPro; IPR005119; LysR_subst.
 DR Pfam; PF00126; HTH_1; 1.
 DR Pfam; PF03466; LysR_substrate; 1.
 DR PRINTS; PR00039; HTHLYSR.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
 KW DNA-binding; Hypothetical protein; Transcription regulation.
 SQ SEQUENCE 307 AA; 33707 MW; D6996094B39047C3 CRC64;

Query Match 100.0%; Score 26; DB 2; Length 307;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5
 Db 230 KHAVE 234

RESULT 20
 Q8ZX44 PRELIMINARY; PRT; 315 AA.
 ID Q8ZX44
 AC Q8ZX44;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Aminopeptidase from family M42.
 GN PAE1474.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;


```

OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
  Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
  aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
DR EMBL; AE009823; AAL63505.1; -;
DR InterPro; IPR001005; Myb DNA binding.
DR InterPro; IPR001917; NHTransf.2.
DR PROSITE; PS00599; AA_TRANSF_CLASS_2; UNKNOWN_1.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 315 AA; 33743 MW; 55485D6D7DD089D0 CRC64;

  Query Match 100.0%; Score 26; DB 17; Length 315;
  Best Local Similarity 100.0%; Pred. No. 1.9e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 301 KHAVE 305

RESULT 21
Q8TNZ4 PRELIMINARY; PRT; 326 AA.
AC Q8TNZ4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein MA2133.
GN MA2133.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=22114;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=1193238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
  FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
  Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
  Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
  Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
  Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A., Leigh J.A.,
  Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K., Springer T.A.,
  Umayam L.A., White O., White R.H., de Macario E.C., Ferry J.G.,
  Jarell K.F., Jing H., Macario A.J.L., Paulsen I., Pritchett M.,
  Sowers K.R., Swanson R.V., Zinder S.H., Lander E., Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
  and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010900; AAM0531.1; -;
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 326 AA; 37316 MW; 055094537DC6F47C CRC64;

  Query Match 100.0%; Score 26; DB 17; Length 326;
  Best Local Similarity 100.0%; Pred. No. 2e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 311 KHAVE 315

RESULT 22

```

```

Q9FRD8
ID Q9FRD8 PRELIMINARY; PRT; 332 AA.
AC Q9FRD8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative peroxidase.
GN OSJNBA0013M12.9
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
  Bowman C.L., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
  Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
  Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBA0013M12 genomic sequence.";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC082644; AAG46142.1; -;
DR HSSP; P22195; 1SCH.
DR InterPro; IPR02016; Peroxidase.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00458; PEROXIDASE.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
KW Peroxidase.
SQ SEQUENCE 332 AA; 35830 MW; 39D62830A5536CEE CRC64;

  Query Match 100.0%; Score 26; DB 10; Length 332;
  Best Local Similarity 100.0%; Pred. No. 2.1e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHAVE 5
Db 113 KHAVE 117

RESULT 23
Q97D39 PRELIMINARY; PRT; 340 AA.
ID Q97D39;
AC Q97D39;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Oligopeptide ABC transporter, ATPase component.
GN CAC3642.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VTM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Neelling J., Breston G., Omelchenko M.V., Makarova K.S., Zeng Q.,
  Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
  Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
  Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
  bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007859; AAK81564.1; -;
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transportr; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Complete proteome.

```


SQ SEQUENCE 340 AA; 37610 MW; 16C446940267BD0E CRC64;
Query Match 100.0%; Score 26; DB 16; Length 340;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5
130 KHAVE 134

RESULT 24

QVNZ6 PRELIMINARY; PRT; 348 AA.
AC QVNZ6;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Putative oxidoreductase.
GN WLEA.
OS Bordetella avium.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=521;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=197N;
RA Spears P.E., Orndorff P.E., Temple L.M., Preston A., Maskell D.J.,
Parkhill J., Mungall K.L., Goodhead I.;
RT "Lipopolysaccharide biosynthesis in Bordetella avium";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ427964; CAD21031.1; -
DR InterPro; IPR00683; GPO_IDH_MoCA.
DR InterPro; IPR004104; GPO_IDH_MoCA_C.
DR Pfam; PF01408; GPO_IDH_MoCA_1.
DR Pfam; PF02894; GPO_IDH_MoCA_C_1.
SQ SEQUENCE 348 AA; 38770 MW; 8950275FC2FFC49A CRC64;

Query Match 100.0%; Score 26; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5
140 KHAVE 144

RESULT 25

Q9KMV6 PRELIMINARY; PRT; 349 AA.
AC Q9KMV6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical protein VCA0212.
GN VCA0212.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; Pubmed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
Gill S.R., Nelson K.E., Read T.D., Tettlin H., Richardson D.,
Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
Zberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
C.M.;
ence of both chromosomes of the cholera pathogen Vibrio
483(2000).

DR EMBL; AE004361; AAF96124.1; -
DR TIGR; VCA0212; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 349 AA; 39620 MW; 55CB09AFOF9AGF2F CRC64;

Query Match 100.0%; Score 26; DB 16; Length 349;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHAVE 5
224 KHAVE 228

Search completed: June 20, 2003, 20:57:04
Job time : 30.3636 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 31.918 seconds
(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-18
Perfect score: 34
Sequence: 1 CVAHC 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : A_Geneseq_101002:*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	5	22	AG65378
2	34	100.0	5	23	AAW47485
3	33	97.1	44	22	AB242931
4	33	97.1	44	22	AB26172
5	33	97.1	44	22	AAW63832
6	33	97.1	44	22	AAW76846
7	33	97.1	44	22	AAW20958
8	33	97.1	44	22	AAW36751
9	33	97.1	44	23	ABG45863
10	33	97.1	86	22	AAW64596

11	33	97.1	94	22	AAW47488
12	33	97.1	134	21	AAW40909
13	33	97.1	134	23	ABP06756
14	33	97.1	166	22	ABG08149
15	33	97.1	166	22	ABW47563
16	31	91.2	25	20	AAW01454
17	31	91.2	47	22	ABW28739
18	31	91.2	47	22	ABW33923
19	31	91.2	47	22	ABW13163
20	31	91.2	47	22	AAW54690
21	31	91.2	47	22	AAW67091
22	31	91.2	47	22	AAW14949
23	31	91.2	47	22	AAW27382
24	31	91.2	47	22	AAW02676
25	31	91.2	47	23	ABG36749
26	31	91.2	58	23	ABP05836
27	31	91.2	68	22	ABW03685
28	31	91.2	96	22	AAW90636
29	31	91.2	107	22	AAW02476
30	31	91.2	230	20	AAW60056
31	31	91.2	258	22	ABG29144
32	31	91.2	297	22	AAW40597
33	31	91.2	298	22	AAW38811
34	31	91.2	298	22	AAW93770
35	31	91.2	449	21	AAW45924
36	31	91.2	482	23	AAU98427
37	31	91.2	483	21	AAW25784
38	31	91.2	483	21	AAW45923
39	31	91.2	485	21	AAW25783
40	31	91.2	485	21	AAW45922
41	31	91.2	485	22	AAW19934
42	31	91.2	500	22	AAW19936
43	31	91.2	503	21	AAW25782
44	31	91.2	1013	22	ABG02686
45	31	91.2	1043	21	AAW53669
46	31	91.2	1043	21	AAW53671
47	31	91.2	1062	21	AAW53674
48	31	91.2	1118	21	AAW53670
49	31	91.2	1118	21	AAW53673
50	31	91.2	3021	22	ABG02194
51	30	88.2	5	22	AAW65379
52	30	88.2	5	23	AAW47486
53	30	88.2	17	18	AAW01831
54	30	88.2	20	18	AAW01834
55	30	88.2	38	22	ABW15131
56	30	88.2	39	19	AAW66455
57	30	88.2	39	21	AAW91754
58	30	88.2	40	11	AAW00720
59	30	88.2	40	14	AAW00780
60	30	88.2	40	14	AAW30562
61	30	88.2	40	19	AAW66454
62	30	88.2	40	19	AAW66435
63	30	88.2	40	21	AAW91734
64	30	88.2	40	21	AAW91753
65	30	88.2	40	23	AAU91037
66	30	88.2	42	9	AAW82390
67	30	88.2	43	16	AAW82841
68	30	88.2	43	16	AAW82842
69	30	88.2	43	18	AAW26000
70	30	88.2	43	23	AAU91058
71	30	88.2	45	15	AAW57129
72	30	88.2	45	21	AAW88018
73	30	88.2	53	23	ABP02851
74	30	88.2	61	22	AAW01668
75	30	88.2	63	21	AAW00377

ALIGNMENTS

RESULT 1
AAG65378
ID AAG65378 standard; peptide; 5 AA.

Human protease and
Human ORFX ORF673
Human ORFX protein
Novel human diapo
Protease PRTS-5.
Secreted protein e
Peptide #1390 enco
Peptide #1429 enco
Protein #1362 enco
Human brain expres
Human bone marrow
Peptide #1383 enco
Peptide #1419 enco
Peptide #1358 enco
Human ORFX protein
Human musculoskele
Human secreted pro
Human polypeptide
Human endometrium
Novel human diapo
Human polypeptide
Human polypeptide
Human protein sequ
Arabidopsis thalia
Cadium-resistance
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Wheat phytochelati
Arabidopsis thalia
Novel human diapo
Mechanical stress
Protein 405 no up
Sequence gif/113642
Mechanical stress
Protein 405 hum se
Novel human diapo
Cyclic control pep
Cyclic control pep
Tenecin peptide fr
Tenecin peptide fr
Novel human diapo
Cationic peptide p
Cationic peptide p
Phormia terranova
Phormia terranova
Spectin - an antib
Cationic peptide p
Cationic peptide s
Cationic peptide p
Transplant media a
Transplant media a
KM-2 antiviral and
Allomycin A antiba
Allomycin B antiba
Tenecin. Tenebrio
Transplant media a
Yeast pro-AMF C-te
Human growth hormo
Human ORFX protein
Human polypeptide
Human secreted pro

XX AC AAG65378;
 XX 30-NOV-2001 (first entry)
 XX
 XX Cyclic peptide with cadherin CAR sequence.
 XX
 XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 KW cytosolic; vulnerable; immunomodulator; vasotropic; neuroprotective;
 KW cerebroprotective; muscular; cyclic.
 XX OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "putative N-terminal acetylation or
 FT alkoxymethylbenzyl" /note= "putative C-terminal amide or ester"
 FT Modified-site 5
 FT
 FT WO200153331-A2.
 XX
 XX 26-JUL-2001.
 XX
 XX 24-JAN-2001; 2001WO-US02508.
 XX
 XX 24-JAN-2000; 2000US-0491078.
 PR
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 PI Hu Z;
 XX WPI; 2001-549899/61.
 XX
 XX Cell adhesion modulating agent used for enhancing delivery of drug to
 PT tumor comprises imidazole compounds -
 PT
 XX Examples; Fig 3A-I; 436pp; English.
 XX
 XX The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cyclic peptide with
 CC cadherin cell adhesion recognition (CAR) sequence.
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 34; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 DB 1 CVAHC 5
 RESULT 2
 AAM47485
 ID AAM47485 standard; peptide; 5 AA.
 XX
 XX AAM47485;
 XX
 DT 12-FEB-2002 (first entry)

XX Cyclic control peptide #5.
 DE
 XX
 KW Cadherin; cytostatic; gynecological; endometriosis;
 KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 KW cyclic.
 XX OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "Linked to residue 5 to form a cyclic peptide,
 FT N-terminal acetyl"
 FT Modified-site 5 /note= "Linked to residue 1 to form a cyclic peptide,
 FT C-terminal amide"
 XX
 XX WO200177146-A2.
 PN
 XX 18-OCT-2001.
 PD
 XX 09-APR-2001; 2001WO-US11669.
 XX
 XX 07-APR-2000; 2000US-0544782.
 PR
 XX (UTMC-) UNIV MCGILL.
 PA
 XX Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 PI WPI; 2002-049129/06.
 DR
 XX Modulating endothelial cell adhesion for inhibiting development of
 PT endometriosis, increasing blood flow to tumor in a mammal, by
 PT contacting cell with a cyclic peptide having cadherin cell adhesion
 PT recognition sequence -
 PT
 XX Disclosure; Fig 3; 139pp; English.
 XX
 XX The present invention relates to a method for modulating endothelial cell
 CC adhesion. The method comprises contacting an endothelial cell with a
 CC modulating peptide comprising a cadherin cell adhesion recognition
 CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 CC for inhibiting angiogenesis, increasing vasopermeability, increasing
 CC blood flow to a tumour, disrupting neovasculature and inhibiting the
 CC development of endometriosis in a mammal. The modulating peptide reduces
 CC unwanted endothelial adhesion occurring between tumour cells, tumour
 CC cells and normal cells, normal cells as a result of surgery, injury,
 CC chemotherapy, disease and inflammation. The present sequence is a peptide
 CC used as a control for illustrating the present invention.
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 34; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 DB 1 CVAHC 5
 RESULT 3
 ABB42931
 ID ABB42931 standard; Peptide; 44 AA.
 XX
 XX ABB42931;
 XX
 XX 04-FEB-2002 (first entry)
 DT
 XX Peptide #10437 encoded by human foetal liver single exon probe.
 DE
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW
 XX


```

XX      SQ      Sequence      44 AA;
      Query Match      97.1%; Score 33; DB 22; Length 44;
      Best Local Similarity 80.0%; Pred. NO. 1.2e+02;
      Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 CVAHC 5
      ||||
      30 CIAHC 34

Db

RESULT 7
AAM20958
ID AAM20958 standard; Protein: 44 AA

```

XX
AC
XX
DT
XX
DE
XX
KW
KW
KW
XX

AAM20958;
12-OCT-2001 (first entry)
Peptide #7392 encoded by probe for measuring cervical gene expression.
Probe; human; microarray; gene expression; cervical epithelial cell;
cervical cancer.

XX	WO200157278-A2.
PN	
XX	
XX	09-AUG-2001.
XX	
XX	30-JAN-2001; 2001WO-US006670.
PF	
XX	
XX	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0234639.
PR	

PR XX
PA XX
PI XX
DR XX
WPI XX

04-OCT-2000; 2000GB-0024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human cervical epithelial cells -
Claim 27; SEQ ID No 25784; 487pp; English.
The present invention relates to human single exon nucleic acid probes
(SENP; see AH10068-AH12845). The present sequence is a peptide encoded

CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 44 AA;

```

Query Match      97.1%; Score 33; DB 22; Length 44;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY      1 CVAHC 5
Db      30 CTAHC 34

```

RESULT 8
AAM36751
ID AAM36751 standard; Protein; 44 AA.
XX AC AAM36751;
XX DT 17-OCT-2001 (first entry)
XX DE Peptide #10788 encoded by probe for measuring placental gene expression.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW Genetic disorder;
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-48897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human placenta -
XX PS Claim 27; SEQ ID No 37020; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENPs;
XX CC see AAI31315-AA15746). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders.
XX SQ Sequence 44 AA;
Query Match 97.1%; Score 33; DB 22; Length 44;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CVAHC 5
Db 30 CIAHC 34
RESULT 9
ABG45863
ID ABG45863 standard; Peptide; 44 AA.
XX AC ABG45863;
XX DT 19-AUG-2002 (first entry)
XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 35528.
XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX KW Chronic obstructive pulmonary disease; interstitial lung disease;
XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberos scleriosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX OS Homo sapiens.
XX PN WO200186003-A2.
XX PD 15-NOV-2001.
XX PF 30-JAN-2001; 2001WO-US00665.
XX PR 04-FEB-2000; 2000US-180312P.
XX PR 26-MAY-2000; 2000US-207456P.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-234687P.
XX PR 27-SEP-2000; 2000US-236359P.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX PT Spatially-addressable set of single exon nucleic acid probes, used to
XX PT measure gene expression in human lung samples -
XX PS Claim 27; SEQ ID No 35528; 634pp; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human lung comprising single exon nucleic acid probes having one of
XX CC 12614 nucleic acid sequences mentioned in the specification, or their
XX CC complements or the 12387 open reading frames derived from the 12614
XX CC probes. Also included are a microarray comprising the novel set of
XX CC probes; the novel set of probes which hybridize at high stringency to a
XX CC nucleic acid expressed in the human lung; measuring gene expression in a
XX CC sample derived from human lung, comprising (a) contacting the array with
XX CC a collection of detectably labeled nucleic acids derived from human lung
XX CC mRNA, and (b) measuring the label detectably bound to each probe of
XX CC the array; identifying exons in a eukaryotic genome, comprising
XX CC (a) algorithmically predicting at least one exon from genomic sequences
XX CC of the eukaryote; and (b) detecting specific hybridisation of detectably
XX CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX CC having a fragment identical to the predicted exon, the probe is included
XX CC in the above mentioned microarray; assigning exons to a single gene,
XX CC comprising (a) identifying exons from genomic sequence by the method
XX CC above and (b) measuring the expression of each of the exons in several
XX CC tissues and/or cell types using hybridisation to a single exon
XX CC microarrays having a probe with the exon, where a common pattern of
XX CC expression of the exons in the tissues and/or cell types indicates that
XX CC the exons should be assigned to a single gene; a peptide comprising one
XX CC of 12011 sequences, mentioned in the specification, or encoded by the
XX CC probes/open reading frames (ORF). The probes are used for gene
XX CC expression analysis, and for identifying exons in a gene, particularly
XX CC using human lung derived mRNA and for the study of lung diseases
XX CC such as asthma, lung cancer, chronic obstructive pulmonary disease
XX CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
XX CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX CC and hyaline membrane disease. The present sequence is a peptide/protein
XX CC encoded by a single exon probe of the invention.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 44 AA;

SQ Query Match 97.1%; Score 33; DB 23; Length 44;

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVAHC 5

DB 30 CIAHC 34;

RESULT 10

AAB64596

ID AAB64596 standard; Protein; 86 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVAHC 5

DB 30 CIAHC 34;

RESULT 10

AAB64596

ID AAB64596 standard; Protein; 86 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVAHC 5

DB 30 CIAHC 34;

RESULT 10

AAB64596

ID AAB64596 standard; Protein; 86 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVAHC 5

DB 30 CIAHC 34;

RESULT 10

AAB64596

ID AAB64596 standard; Protein; 86 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVAHC 5

DB 30 CIAHC 34;

RESULT 10

AAB64596

ID AAB64596 standard; Protein; 86 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

CC atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PPM
 CC polynucleotide sequences can be used in somatic or germline gene therapy
 CC and in diagnosis of diseases. They can also be used in generating
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequences and in molecular biology techniques.

XX
 SQ Sequence 94 AA;
 Query Match 97.1%; Score 33; DB 22; Length 94;
 Best Local Similarity 80.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
 Db 21 CIAHC 25

RESULT 12
 AAB40909
 ID AAB40909 standard; Protein; 134 AA.
 AC AAB40909;
 XX
 DT 08-FEB-2001 (first entry)
 DE Human ORFX ORF673 polypeptide sequence SEQ ID NO:1346.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antichyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC75118.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 1152; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;

CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antichyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX
 SQ Sequence 134 AA;
 Query Match 97.1%; Score 33; DB 21; Length 134;
 Best Local Similarity 80.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
 Db 130 CIAHC 134

RESULT 13
 AAB06756
 ID AAB06756 standard; Protein; 134 AA.
 AC AAB06756;
 XX
 DT 24-JUN-2002 (first entry)
 DE Human ORFX protein sequence SEQ ID NO:13494.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

OS Homo sapiens.
 XX
 PN WO200192523-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US10836.
 XX
 PR 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach MD;
 XX
 DR WPI; 2002-106308/14.
 DR N-PSDB; AAB22508.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease; neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 PS Disclosure; SEQ ID 13494; 1037pp; English.
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see table 1

in the specification). ABN15762 to ABN27252 encode the human ORFX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for treating or preventing a pathology associated with an ORFX-associated disorder in humans, and in the manufacture of a medicament for treating a syndrome associated with ORFX-associated disorder. ORFX polynucleotide sequences can be used in gene therapy. ORFX sequences can be used in the treatment of cancer, hyperproliferative disorders, cirrhosis of liver, psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORFX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage.

N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match 97.1%; Score 33; DB 23; Length 134;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|
Db 130 CIAHC 134

RESULT 14
ABG08149
ID ABG08149 standard; Protein; 166 AA.

AC ABG08149;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8140.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC,

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS72336.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity

XX Claim 20; SEQ ID No 38508; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 166 AA;

Query Match 97.1%; Score 33; DB 22; Length 166;

Best Local Similarity 80.0%; Pred. No. 3.7e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5

|:|
Db 21 CIAHC 25

RESULT 15

AAAB47563

ID AAB47563 standard; Protein; 166 AA.

XX AAB47563;

XX 13-DEC-2001 (first entry)

XX Protease PRTS-5.

XX Human; protease; PRTS; gastrointestinal disorder; peptic esophagitis;
KW indigestion; gastritis; cardiovascular disorder; reproductive disorder;
KW hypertensive heart disease; myocardial infarction; autoimmune disorder;
KW inflammatory disorder; rheumatoid arthritis; cell proliferative disorder;
KW arteriosclerosis; cancer; epithelial disorder; eczema; endometriosis;
KW neurological disorder; Huntington's disease.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 9 /label= Potential phosphorylation site

XX Modified-site 55 /label= Potential glycosylation site

XX Modified-site 73 /label= Potential phosphorylation site

XX Modified-site 97 /label= Potential phosphorylation site

XX Modified-site 122 /label= Potential glycosylation site

XX WO200171004-A2.

XX 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US08441.

XX 17-MAR-2000; 2000US-190708P.

XX 30-MAR-2000; 2000US-193182P.

XX 14-APR-2000; 2000US-197086P.

PR 20-APR-2000; 2000US-199022P.
 PR 28-APR-2000; 2000US-200227P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Yue H, Lu DAM, Pollicky JL, Delegeane AM, Tribouley CM, Khan FA;
 PI Au-Young J, Bandman O, Lal P, Borowsky ML, Gandhi AR, Hillman JL;
 PI Tang YT, Burford N, Baughn MR, Nguyen DB, Yao MG, Wallia NK, He A;
 PI Hafalia A, Lu Y, Patterson C;
 XX WPI; 2001-611509/70.
 DR N-PSDB; AAH43516.
 XX New polypeptide for treating gastrointestinal, cardiovascular and
 PT autoimmune disorders, comprises novel human proteases (PRTS) and,
 PT polynucleotides -
 XX Claim 1; Page 115; 129pp; English.
 XX The sequences given in AAB47559-69 show novel human proteases PRTS-1
 CC to -11. These proteases can be administered in pharmaceutical compounds
 CC for the treatment of diseases or conditions associated with decreased
 CC expression of functional protease. Disorders which can be treated
 CC include gastrointestinal (peptic esophagitis, indigestion and
 CC gastritis), cardiovascular (hypertensive heart disease, myocardial
 CC infarction), autoimmune/inflammatory (rheumatoid arthritis), cell
 CC proliferative (arteriosclerosis, cancers), epithelial (eczema),
 CC neurological (Huntington's disease), and reproductive (endometriosis)
 CC disorders. This protein has homology to the S. pombe putative
 CC mitochondrial membrane protease subunit 2.
 XX Sequence 166 AA;
 SQ Query Match 97.1%; Score 33; DB 22; Length 166;
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 Db |||||
 21 CIAHC 25
 RESULT 16
 AA01454
 ID AAY01454 standard; Protein; 25 AA.
 XX AAY01454;
 AC AAY01454;
 XX 18-MAY-1999 (first entry)
 DT Secreted protein encoded by gene 64 clone HNECF34.
 DE Human; secreted protein; gene therapy; protein therapy; cancer; weight;
 KW tumour; chromosome mapping; forensic; haematological disease; allergy;
 KW inflammation; cell proliferation; viral infection; wound healing;
 KW modulation; appetite; behaviour; food additive; preservative.
 XX Homo sapiens.
 OS WO9903990-A1.
 XX 28-JAN-1999.
 PD 15-JUL-1998; 98WO-US14613.
 XX 18-AUG-1997; 97US-0056361.
 PR 16-JUL-1997; 97US-0052661.
 PR 16-JUL-1997; 97US-0052870.
 PR 16-JUL-1997; 97US-0052871.
 PR 16-JUL-1997; 97US-0052872.
 PR 16-JUL-1997; 97US-0052873.
 PR 16-JUL-1997; 97US-0052874.
 PR 16-JUL-1997; 97US-0052875.

PR 22-JUL-1997; 97US-0053440.
 PR 22-JUL-1997; 97US-0053441.
 PR 22-JUL-1997; 97US-0053442.
 PR 18-AUG-1997; 97US-0055683.
 PR 18-AUG-1997; 97US-0055724.
 PR 18-AUG-1997; 97US-0055725.
 PR 18-AUG-1997; 97US-0055726.
 PR 18-AUG-1997; 97US-0055946.
 PR 18-AUG-1997; 97US-0055952.
 PR 18-AUG-1997; 97US-0055985.
 PR 18-AUG-1997; 97US-0055989.
 PR 18-AUG-1997; 97US-0056359.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Duan R, Peng P, Ferrie AM, Florence KA, Fouad J;
 PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE;
 PI Yu G;
 XX WPI; 1999-132234/11.
 DR N-PSDB; AAX22282.
 XX New nucleic acids encoding secreted human proteins - potentially
 PT useful for treating and diagnosing diseases and identifying specific
 PT binding agents
 XX Claim 11; Page 238; 251pp; English.
 PS The invention relates to nucleic acid sequences (AAX22211 to AAX22282)
 CC encoding human secreted proteins (AAY01383 to AAY01454). The secreted
 CC protein gene sequences are deposited with the ATCC under deposit number
 CC ATCC 209139, 209139 or 209141. Host cells containing vectors comprising
 CC the nucleic acid sequences are used for the recombinant expression of
 CC the secreted proteins. The polynucleotide and amino acid sequences are
 CC useful for preventing, treating or ameliorating medical conditions e.g.
 CC by protein or gene therapy. Pathological conditions can be also diagnosed
 CC by determining the amount of the new polypeptides in a sample or by the
 CC presence of mutations in the new polynucleotides. The nucleic acid
 CC sequences, or its fragments, are useful for chromosome identification
 CC and mapping; as antisense and triplex-forming therapeutics; in gene
 CC therapy; for (forensic) identification of individuals; as molecular
 CC weight markers; to identify related sequences or specific mRNA; in
 CC preparation of oligomers and to raise anti-DNA antibodies. Antibodies are
 CC useful as immunoassay reagents (including for in vivo imaging) and
 CC therapeutically to inhibit or activate particular polypeptides. A very
 CC wide range of disorders may be treated with the polynucleotide and
 CC polypeptide sequences, e.g. autoimmune or haematological diseases,
 CC allergy, inflammation, cancer or other forms of cell proliferation, viral
 CC or other infections. The sequences may also be useful in wound healing,
 CC to modulate differentiation of embryonic stem cells, to modulate weight,
 CC appetite, behaviour etc. and as food additive or preservative. The
 CC present sequence represents a human secreted protein (see descriptor
 CC line for gene number and clone identification).
 XX Sequence 25 AA;
 SQ Query Match 91.2%; Score 31; DB 20; Length 25;
 Best Local Similarity 80.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 Db |||||
 13 CVSHC 17
 RESULT 17
 ABB28739
 ID ABB28739 standard; Peptide; 47 AA.
 XX ABB28739;
 AC ABB28739;
 XX 01-FEB-2002 (first entry)
 DT

DE Peptide #1390 encoded by breast cell single exon nucleic acid probe.
XX Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX Homo sapiens.
XX WO200157271-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00662.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX Claim 27; SEQ ID NO 11707; 327pp + sequence listing; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 47 AA;
Query Match 91.2%; Score 31; DB 22; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1'CVAVHC 5
|:|:|
Db 13 CLAVHC 17
RESULT 18
ABB33923
ID ABB33923 standard; Peptide; 47 AA.
XX AC ABB33923;
XX 04-FEB-2002 (first entry)
XX Peptide #1429 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00669.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX Claim 27; SEQ ID NO 26558; 639pp + sequence listing; English.
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 47 AA;
Query Match 91.2%; Score 31; DB 22; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1'CVAVHC 5
|:|:|
Db 13 CLAVHC 17
RESULT 19
ABB19363
ID ABB19363 standard; Protein; 47 AA.
XX AC ABB19363;
XX 23-JAN-2002 (first entry)
XX Protein #1362 encoded by probe for measuring heart cell gene expression.
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX Homo sapiens.
XX WO200157274-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -

XX Claim 15; SEQ ID No 21133; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting, of the
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

XX Sequence 47 AA;

Query Match 91.2%; Score 31; DB 22; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|:|
Db 13 CLAHC 17

RESULT 20

AAAM54690
ID AAAM54690 standard; Protein; 47 AA.

XX

AC AAAM54690;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 26795.

XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -

XX Example 4; SEQ ID NO: 26795; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.

XX Sequence 47 AA;

Query Match 91.2%; Score 31; DB 22; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|:|
Db 13 CLAHC 17

RESULT 21

AAAM67091
ID AAAM67091 standard; Protein; 47 AA.

XX

AC AAAM67091;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 27397.
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 27397; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.

XX SQ Sequence 47 AA;

Query Match 91.2%; Score 31; DB 22; Length 47;

Best Local Similarity 80.0%; Pred. No. 2.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|
Db 13 CLAHC 17

RESULT 22

AAM14949

ID AAM14949 standard; Protein; 47 AA.

XX AC AAM14949;

DT 12-OCT-2001 (first entry)

DE Peptide #1383 encoded by probe for measuring cervical gene expression.

KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW Cervical cancer.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00670.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human cervical epithelial cells -

XX PS Claim 27; SEQ ID No 19775; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded
XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX CC can be used to produce a single exon microarray, which can be used for
XX CC measuring human gene expression in a sample derived from human cervical
XX CC epithelial cells. By measuring gene expression, the probes are therefore
XX CC useful in grading and/or staging of diseases of the cervix, notably
XX CC cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 47 AA;

Query Match 91.2%; Score 31; DB 22; Length 47;

Best Local Similarity 80.0%; Pred. No. 2.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|
Db 13 CLAHC 17

RESULT 24

AAM02676

ID AAM02676 standard; Protein; 47 AA.

XX AC AAM02676;

XX DT 09-OCT-2001 (first entry)

XX DE Peptide #1358 encoded by probe for measuring breast gene expression.

QY 1 CVAHC 5
|:|
Db 13 CLAHC 17

RESULT 23

AAM27382

ID AAM27382 standard; Protein; 47 AA.

XX AC AAM27382;

DT 17-OCT-2001 (first entry)

DE Peptide #1419 encoded by probe for measuring placental gene expression.

KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human placenta -

XX PS Claim 27; SEQ ID No 27651; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AAI13115-AA157546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders.

XX SQ Sequence 47 AA;

Query Match 91.2%; Score 31; DB 22; Length 47;

Best Local Similarity 80.0%; Pred. No. 2.6e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|
Db 13 CLAHC 17

RESULT 24

AAM02676

ID AAM02676 standard; Protein; 47 AA.

XX AC AAM02676;

XX DT 09-OCT-2001 (first entry)

XX DE Peptide #1358 encoded by probe for measuring breast gene expression.

KW Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX

OS Homo sapiens.
XX

PN WO200157270-A2.
XX

XX 09-AUG-2001.
XX

XX 29-JAN-2001; 2001WO-US000661.
XX

XX 04-FEB-2000; 2000US-0180312.
XX

XX 26-MAY-2000; 2000US-0207456.
XX

XX 30-JUN-2000; 2000US-0608408.
XX

XX 03-AUG-2000; 2000US-0632366.
XX

XX 21-SEP-2000; 2000US-0234687.
XX

XX 27-SEP-2000; 2000US-0236359.
XX

XX 04-OCT-2000; 2000GB-0024263.
XX

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX

XX WPI; 2001-476286/51.
XX

XX Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX

XX Claim 27; SEQ ID No 11416; 322pp; English.
XX

XX The present invention relates to novel single exon nucleic acid probes
XX (see AAT00010-AAL10067). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for measuring human gene expression in
XX a human breast sample, where the probe hybridises at high stringency to a
XX nucleic acid expressed in the human breast. The probes are useful for
XX predicting, diagnosing, grading, staging, monitoring and prognosing
XX diseases of the human breast, particularly those diseases with polygenic
XX aetiology. The diseases include: breast cancer, disorders of development,
XX inflammatory diseases of the breast, fibrocystic changes, proliferative
XX breast disease and non-carcinoma tumours.
XX

XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX

XX Sequence 47 AA;
XX

XX Query Match 91.2%; Score 31; DB 22; Length 47;
XX Best Local Similarity 80.0%; Pred. No. 2.6e+02;
XX Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX

XX 1 CUNAC 5
XX

XX 13 CLAHC 17
XX

XX RESULT 25
XX

XX ABG36749
XX

XX ID ABG36749 standard; Peptide; 47 AA.
XX

XX AC ABG36749;
XX

XX 19-AUG-2002 (first entry)
XX

XX Human peptide encoded by genome-derived single exon probe SEQ ID 26414.
XX

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX

KW hyaline membrane disease.
XX

OS Homo sapiens.
XX

PN WO200186003-A2.
XX

XX 15-NOV-2001.
XX

XX 30-JAN-2001; 2001WO-US000665.
XX

XX 04-FEB-2000; 2000US-180312P.
XX

XX 26-MAY-2000; 2000US-207456P.
XX

XX 30-JUN-2000; 2000US-0608408.
XX

XX 03-AUG-2000; 2000US-0632366.
XX

XX 21-SEP-2000; 2000US-234687P.
XX

XX 27-SEP-2000; 2000US-236359P.
XX

XX 04-OCT-2000; 2000GB-0024263.
XX

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX

XX WPI; 2002-114183/15.
XX

XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX

XX Claim 27; SEQ ID No 26414; 634pp; English.
XX

XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a peptide/protein
XX encoded by a single exon probe of the invention.
XX

XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX

XX Sequence 47 AA;
XX

XX Query Match 91.2%; Score 31; DB 23; Length 47;
XX

Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
|:|
Db 13 CLAHC 17

Search completed: June 20, 2003, 20:51:59
Job time : 33.9318 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 10.5682 Seconds
(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-18

Perfect score: 34

Sequence: 1 CVAHC 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

1: /cgn2_6/prodata/1/aaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/aaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/aaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/aaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/aaa/CTUS_COMB.pep.*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	5	4	US-09-115-395-70
2	34	100.0	5	4	US-09-250-059-18
3	34	100.0	5	4	US-09-248-074-18
4	34	100.0	5	4	US-09-357-717-15
5	34	100.0	5	4	US-09-458-870-18
6	30	88.2	5	4	US-09-115-395-71
7	30	88.2	5	4	US-09-250-059-19
8	30	88.2	5	4	US-09-248-074-19
9	30	88.2	5	4	US-09-458-870-19
10	30	88.2	40	1	US-08-385-375-16
11	30	88.2	40	1	US-08-385-375-39
12	30	88.2	210	4	US-09-171-461-34
13	30	88.2	288	1	US-08-368-852-15
14	30	88.2	288	1	US-08-525-940-15
15	30	88.2	288	2	US-08-976-838-15
16	30	88.2	327	1	US-08-080-386-2
17	30	88.2	327	2	US-08-390-000A-2
18	30	88.2	327	4	US-09-583-491-10
19	30	88.2	328	1	US-08-423-691-2
20	30	88.2	328	1	US-08-080-386-4
21	30	88.2	328	2	US-08-390-000A-4
22	30	88.2	328	3	US-08-867-260-2
23	30	88.2	328	4	US-09-501-192-9
24	30	88.2	359	3	US-08-586-165-4
25	30	88.2	615	2	US-08-663-566A-17
26	30	88.2	615	2	US-08-023-610-17
27	30	88.2	615	2	US-08-288-065A-17

28	30	88.2	615	2	US-08-362-240A-17
29	30	88.2	615	5	PCT-US95-10245-17
30	30	88.2	799	2	US-08-525-940-23
31	30	88.2	799	2	US-08-976-838-23
32	30	88.2	881	2	US-08-525-940-21
33	30	88.2	881	2	US-08-976-838-21
34	30	88.2	915	2	US-08-818-070-2
35	30	88.2	915	2	US-08-723-585-2
36	30	88.2	915	2	US-08-525-940-18
37	30	88.2	915	2	US-08-976-838-18
38	30	88.2	915	4	US-09-214-555B-2
39	30	88.2	915	4	US-09-214-555B-7
40	30	88.2	1162	2	US-08-663-566A-15
41	30	88.2	1162	2	US-08-023-610-15
42	30	88.2	1162	2	US-08-288-065A-15
43	30	88.2	1162	2	US-08-362-240A-15
44	30	88.2	1162	5	PCT-US95-10245-15
45	30	88.2	1342	1	US-07-978-895-4
46	30	88.2	1342	2	US-08-484-438-9
47	30	88.2	1342	2	US-08-473-119-4
48	30	88.2	1342	2	US-08-475-352-4
49	30	88.2	1343	6	5183884-4
50	30	88.2	4544	1	US-08-469-486-52
51	30	88.2	4544	2	US-08-469-658-52
52	29	85.3	63	4	US-09-149-476-445
53	29	85.3	383	4	US-08-857-076-105
54	29	85.3	709	4	US-08-753-247-6
55	29	85.3	713	4	US-08-753-247-9
56	29	85.3	719	4	US-08-753-247-12
57	29	85.3	794	1	US-07-885-972A-2
58	29	85.3	794	1	US-07-885-972A-4
59	29	85.3	794	2	US-08-865-203-2
60	29	85.3	794	2	US-08-745-880-2
61	29	85.3	794	2	US-08-745-880-4
62	29	85.3	794	2	US-08-480-382-2
63	29	85.3	794	2	US-08-480-382-4
64	29	85.3	794	2	US-07-849-420-2
65	29	85.3	794	4	US-09-253-854-2
66	29	85.3	794	4	US-08-953-424-2
67	28	82.4	67	1	US-08-435-040-3
68	28	82.4	67	4	US-09-020-216-3
69	28	82.4	145	2	US-08-788-584-1
70	28	82.4	149	2	US-08-788-584-3
71	28	82.4	311	3	US-08-946-914-4
72	28	82.4	311	4	US-09-656-450-4
73	28	82.4	900	4	US-08-890-865A-4
74	28	82.4	992	4	US-08-890-865A-1
75	28	82.4	1956	4	US-08-843-417-10

ALIGNMENTS

RESULT 1
US-09-115-395-70
; Sequence 70, Application US/09115395A
; Patent No. 6207639
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C3
; CURRENT APPLICATION NUMBER: US/09/115,395A
; CURRENT FILING DATE: 1998-07-14
; EARLIER APPLICATION NUMBER: 08/996,679
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/893,534
; EARLIER FILING DATE: 1997-07-11
; EARLIER APPLICATION NUMBER: 60/021,612
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70

; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
; OTHER INFORMATION: Synthesis
US-09-115-395-70

Query Match 100.0%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | | |
Db 1 CVAHC 5

RESULT 2
US-09-250-059-18
; Sequence 18, Application US/09250059
; Patent No. 633307
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C6
; CURRENT APPLICATION NUMBER: US/09/250,059
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-250-059-18

Query Match 100.0%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | | |
Db 1 CVAHC 5

RESULT 3
US-09-248-074-18
; Sequence 18, Application US/09248074
; Patent No. 6346512
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C5
; CURRENT APPLICATION NUMBER: US/09/248,074
; CURRENT FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic

; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-248-074-18

Query Match 100.0%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | | |
Db 1 CVAHC 5

RESULT 4
US-09-357-717-15
; Sequence 15, Application US/09357717
; Patent No. 6417325
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-357-717-15

Query Match 100.0%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels* 0; Gaps 0;

Qy 1 CVAHC 5
| | | | |
Db 1 CVAHC 5

RESULT 5
US-09-458-870-18
; Sequence 18, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 5

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-458-870-18

Query Match 100.0%; Score 34; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CVAHC 5
|||
Db 1 CVAHC 5

RESULT 6
US-09-115-395-71
Sequence 71, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115,395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996,679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893,534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 71
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis
US-09-115-395-71

Query Match 88.2%; Score 30; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 2e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 1;

QY 1 CVAHC 5
|||
Db 1 CVGHC 5

RESULT 7
US-09-250-059-19
Sequence 19, Application US/09250059
Patent No. 6333307
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250,059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19

LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: control peptide
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-250-059-19

Query Match 88.2%; Score 30; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 2e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 1;

QY 1 CVAHC 5
|||
Db 1 CVGHC 5

RESULT 8
US-09-248-074-19
Sequence 19, Application US/09248074
Patent No. 6346512
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
FILE REFERENCE: 100086.401C5
CURRENT APPLICATION NUMBER: US/09/248,074
CURRENT FILING DATE: 1999-02-10
NUMBER OF SEQ ID NOS: 81
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: control peptide
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-248-074-19

Query Match 88.2%; Score 30; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 2e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 1;

QY 1 CVAHC 5
|||
Db 1 CVGHC 5

RESULT 9
US-09-458-870-19
Sequence 19, Application US/09458870
Patent No. 6465427
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
FILE REFERENCE: 100086.401C8
CURRENT APPLICATION NUMBER: US/09/458,870
CURRENT FILING DATE: 1999-12-10
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: control peptide
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-19

Query Match 88.2%; Score 30; DB 4; Length 5;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
|||
Db 1 CVGHC 5

RESULT 10
US-08-385-375-16
; Sequence 16, Application US/08385375
; Patent No. 5631144
; GENERAL INFORMATION:
; APPLICANT: LEMOINE, Yves
; APPLICANT: NGUYEN, Martine
; APPLICANT: ACHSTETTER, Tilman
; APPLICANT: REICHART, Jean-Marc
; TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A
; TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF
; TITLE OF INVENTION: MATURE PROTEINS BY RECOMBINANT YEASTS, EXPRESSION
; TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS FOR
; TITLE OF INVENTION: THE PREPARATION OF PROTEINS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,375
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,356
; FILING DATE: 04-JAN-1994
; APPLICATION NUMBER: FR 89/05687
; FILING DATE: 28-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR90/00306
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-385-375-16

Query Match 88.2%; Score 30; DB 1; Length 40;
Best Local Similarity 80.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
|||
Db 16 CAHHC 20

RESULT 11
US-08-385-375-39
; Sequence 39, Application US/08385375
; Patent No. 5631144
; GENERAL INFORMATION:
; APPLICANT: LEMOINE, Yves
; APPLICANT: NGUYEN, Martine
; APPLICANT: ACHSTETTER, Tilman
; APPLICANT: REICHART, Jean-Marc
; TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A
; TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF
; TITLE OF INVENTION: MATURE PROTEINS BY RECOMBINANT YEASTS, EXPRESSION
; TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS F
; TITLE OF INVENTION: THE PREPARATION OF PROTEINS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,375
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,356
; FILING DATE: 04-JAN-1994
; APPLICATION NUMBER: FR 89/05687
; FILING DATE: 28-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR90/00306
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-385-375-39

Query Match 88.2%; Score 30; DB 1; Length 40;
Best Local Similarity 80.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
| | | |
Db 16 CTAHC 20

RESULT 12

US-09-171-461-34
; Sequence 34, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 210
; TYPE: PRT
; ORGANISM: CELO VIRUS
; FEATURE:
; OTHER INFORMATION: Position: 4462..5094/note=ORF12
US-09-171-461-34

Query Match 88.2%; Score 30; DB 4; Length 210;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
| | | |
Db 86 CTAHC 90

RESULT 13

US-08-368-852-15
; Sequence 15, Application US/08368852
; Patent No. 5691183
; GENERAL INFORMATION:
; APPLICANT: Franzusoff, Alex
; APPLICANT: Miranda, Luis R.
; TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
; TITLE OF INVENTION: ENCODING SAID PROTEASES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/368,852
; FILING DATE: 05-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: 2848-11
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223

; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-368-852-15

Query Match 88.2%; Score 30; DB 1; Length 288;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
| | | |
Db 95 CVTHC 99

RESULT 14

US-08-525-940-15
; Sequence 15, Application US/08525940
; Patent No. 5866351
; GENERAL INFORMATION:
; APPLICANT: Franzusoff, Alex
; APPLICANT: Miranda, Luis R.
; APPLICANT: Wolf, Joseph R.
; TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
; TITLE OF INVENTION: ENCODING SAID PROTEASES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,940
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/368,852
; FILING DATE: 01-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/088,322
; FILING DATE: 07-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2848-11-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-525-940-15

Query Match 88.2%; Score 30; DB 2; Length 288;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
| | | |
Db 95 CVTHC 99

RESULT 15
US-08-976-838-15
; Sequence 15, Application US/08976838
; Patent No. 5981259
; GENERAL INFORMATION:
; APPLICANT: Fransusoff, Alex
; TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID
; MOLECULES
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/976.838
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2848-11-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-976-838-15
; Query Match 88.2%; Score 30; DB 2; Length 288;
; Best Local Similarity 80.0%; Pred. No. 2.5e+02;
; Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CVAHC 5
Db 95 CVTHC 99
RESULT 16
US-08-080-386-2
; Sequence 2, Application US/08080386
; Patent No. 5750366
; GENERAL INFORMATION:
; APPLICANT: Sealton, Stuart C.
; TITLE OF INVENTION: Cloning and Expression of
; Gonadotropin-Releasing Hormone Receptor
; MOLECULES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/080,386

; FILING DATE: 19930621
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6923-035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-080-386-2
; Query Match 88.2%; Score 30; DB 1; Length 327;
; Best Local Similarity 80.0%; Pred. No. 2.8e+02;
; Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CVAHC 5
Db 195 CVTHC 199
RESULT 17
US-08-390-000A-2
; Sequence 2, Application US/08390000A
; Patent No. 5985583
; GENERAL INFORMATION:
; APPLICANT: Sealton, Stuart C.
; TITLE OF INVENTION: Cloning and Expression of
; Gonadotropin-Releasing Hormone Receptor
; MOLECULES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/390,000A
; FILING DATE: 17-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6923-052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-390-000A-2
; Query Match 88.2%; Score 30; DB 2; Length 327;
; Best Local Similarity 80.0%; Pred. No. 2.8e+02;
; Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CVAHC 5

Db 195 CVTHC 199

RESULT 18

US-09-583-492-10

; Sequence 10, Application US/09583492

; Patent No. 6407222

; GENERAL INFORMATION:

; APPLICANT: Cui, Jisong

; APPLICANT: Lo, Jane-Ling

; APPLICANT: Mount, George R.

; TITLE OF INVENTION: DOG GONADOTROPIN RELEASING HORMONE

; FILE REFERENCE: 20140

; CURRENT APPLICATION NUMBER: US/09/583,492

; CURRENT FILING DATE: 2000-05-31

; EARLIER APPLICATION NUMBER: 60/138,448

; EARLIER FILING DATE: 2000-06-10

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 327

; TYPE: PRT

; ORGANISM: canis

US-09-583-492-10

Query Match

88.2%; Score 30; DB 4; Length 327;

Best Local Similarity 80.0%; Pred. No. 2.8e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

Db 195 CVTHC 199

RESULT 19

US-08-423-691-2

; Sequence 2, Application US/08423691

; Patent No. 5677184

; GENERAL INFORMATION:

; APPLICANT: ONDA, Haruo

; APPLICANT: OHKUBO, Shoichi

; APPLICANT: HINUMA, Shuji

; APPLICANT: SAWADA, Hidekazu

; TITLE OF INVENTION: RECEPTOR EXPRESSION CELLS AND USE

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/423,691

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: RESNICK, David S.

REGISTRATION NUMBER: 34235

REFERENCE/DOCKET NUMBER: 45390

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 523-3400

TELEFAX: (617) 523-6440

TELEX: 200291 STRE UR

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 328 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-423-691-2

Query Match

88.2%; Score 30; DB 1; Length 328;

Best Local Similarity 80.0%; Pred. No. 2.8e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

Db 196 CVTHC 200

RESULT 20

US-08-080-386-4

; Sequence 4, Application US/08080386

; Patent No. 5750366

; GENERAL INFORMATION:

; APPLICANT: Sealton, Stuart C.

; TITLE OF INVENTION: Cloning and Expression of

; TITLE OF INVENTION: Gonadotropin-Releasing Hormone Receptor

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/080,386

FILING DATE: 19930621

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mirock, S. Lealie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 6923-035

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 328 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-080-386-4

Query Match

88.2%; Score 30; DB 1; Length 328;

Best Local Similarity 80.0%; Pred. No. 2.8e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

Db 196 CVTHC 200

RESULT 21

US-08-390-000A-4

; Sequence 4, Application US/08390000A

; Patent No. 5985583

; GENERAL INFORMATION:

; APPLICANT: Sealton, Stuart C.

;; TITLE OF INVENTION: Cloning and Expression of
;; TITLE OF INVENTION: Gonadotropin-Releasing Hormone Receptor
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA: US/08/390,000A
;; APPLICATION NUMBER: US/08/390,000A
;; FILING DATE: 17-FEB-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mierock, S. Leslie
;; REGISTRATION NUMBER: 18,872
;; REFERENCE/DOCKET NUMBER: 6923-052
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212 790-9090
;; TELEFAX: 212 869-8864/9741
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 328 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
US-08-390-000A-4

Query Match 88.2%; Score 30; DB 2; Length 328;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 196 CVTHC 200

RESULT 22
US-08-867-260-2
; Sequence 2, Application US/08867260
; Patent No. 6117645
; GENERAL INFORMATION:
; APPLICANT: ONDA, Haruo
; APPLICANT: OHKUBO, Shoichi
; APPLICANT: HINURA, Shuji
; APPLICANT: SAWADA, Hidekazu
; TITLE OF INVENTION: RECEPTOR EXPRESSION CELLS AND USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,260
; FILING DATE:
; CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/423,691
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: RESNICK, David S
;; REGISTRATION NUMBER: 34235
;; REFERENCE/DOCKET NUMBER: 45390
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617)523-3400
;; TELEFAX: (617)523-6440
;; TELEX: 200291 STRE UR
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 328 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-867-260-2

Query Match 88.2%; Score 30; DB 3; Length 328;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 196 CVTHC 200

RESULT 23
US-09-501-192-9
; Sequence 9, Application US/09501192
; Patent No. 6380375
; GENERAL INFORMATION:
; APPLICANT: Cui, Jisong
; APPLICANT: Lo, Jane-Ling
; APPLICANT: Mount, George R.
; TITLE OF INVENTION: MONKEY GONADOTROPIN RELEASING HORMONE
; TITLE OF INVENTION: RECEPTOR
; FILE REFERENCE: 20139V
; CURRENT APPLICATION NUMBER: US/09/501,192
; CURRENT FILING DATE: 2000-02-10
; EARLIER APPLICATION NUMBER: 60/138,135
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/121,780
; EARLIER FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-501-192-9

Query Match 88.2%; Score 30; DB 4; Length 328;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 196 CVTHC 200

RESULT 24
US-08-586-165-4
; Sequence 4, Application US/08586165
; Patent No. 6054298
; GENERAL INFORMATION:
; APPLICANT: Laufer, Edward M.
; APPLICANT: Orozco, Olivia E.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Fringe Proteins and Pattern Formation
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: Two Militia Drive
 CITY: Lexington
 STATE: Massachusetts
 COUNTRY: US
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/586,165
 FILING DATE: 16-JAN-1996
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: HU95-05
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 861-6240
 TELEFAX: (617) 861-9540
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 359 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..118
 US-08-586-165-4

Query Match 88.2%; Score 30; DB 3; Length 359;
 Best Local Similarity 80.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 Db 289 CTAHC 293

RESULT 25
 US-08-663-566A-17
 ; Sequence 17, Application US/08663566A
 ; Patent No. 5853733
 ; GENERAL INFORMATION:
 ; APPLICANT: Cochran, Mark D
 ; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys
 ; TITLE OF INVENTION: and Uses Thereof
 ; NUMBER OF SEQUENCES: 56
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: John P. White
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/663,566A
 ; FILING DATE: June 13, 1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P
 ; REGISTRATION NUMBER: 28,678
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)278-0400
 ; TELEFAX: (212)391-0526
 ; TELEX: 422523
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 615 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-663-566A-17
 ;
 ; Query Match 88.2%; Score 30; DB 2; Length 615;
 ; Best Local Similarity 80.0%; Pred. No. 5e+02;
 ; Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 ;
 ; Qy 1 CVAHC 5
 ; Db 98 CTAHC 102
 ;
 ; Search completed: June 20, 2003, 21:00:33
 ; Job time : 12.5682 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:52:19 ; Search time 12.8409 Seconds
(without alignments)
42.134 Million cell updates/sec

Title: US-10-105-008-18

Perfect score: 34

Sequence: 1 CVAHC 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	5	9	US-09-769-145-18
2	34	100.0	5	9	US-10-058-821-15
3	33	97.1	44	10	US-09-864-761-41470
4	33	97.1	166	12	US-10-046-643-2
5	31	91.2	24	9	US-09-776-724A-154
6	31	91.2	47	10	US-09-864-761-34661
7	31	91.2	68	10	US-09-764-877-1632
8	31	91.2	96	10	US-09-800-729-179
9	30	88.2	5	9	US-09-769-145-19
10	30	88.2	39	10	US-09-030-619-218
11	30	88.2	40	10	US-09-030-619-198
12	30	88.2	40	10	US-09-030-619-217
13	30	88.2	40	10	US-09-917-340-73
14	30	88.2	43	10	US-09-917-340-94
15	30	88.2	106	9	US-10-196-437A-3
16	30	88.2	129	9	US-10-196-437A-2
17	30	88.2	210	10	US-09-970-711-34
18	30	88.2	906	10	US-09-815-242-10567
19	30	88.2	1342	9	US-10-172-620-16

RESULT 1

US-09-769-145-18
; Sequence 18, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Anmar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoameng

ALIGNMENTS

20	29	85.3	42	9	US-09-764-891-4427	Sequence 4427, Ap
21	29	85.3	47	9	US-10-178-213-378	Sequence 378, App
22	29	85.3	52	9	US-10-178-213-288	Sequence 288, App
23	29	85.3	63	9	US-09-809-391-445	Sequence 445, App
24	29	85.3	73	9	US-10-091-572-409	Sequence 409, App
25	29	85.3	75	9	US-10-106-698-6006	Sequence 6006, Ap
26	29	85.3	78	9	US-10-178-213-377	Sequence 377, App
27	29	85.3	81	10	US-09-764-877-1340	Sequence 1340, Ap
28	29	85.3	82	9	US-10-178-213-287	Sequence 287, App
29	29	85.3	86	9	US-09-764-891-2731	Sequence 2731, Ap
30	29	85.3	96	10	US-09-864-761-43066	Sequence 43066, A
31	29	85.3	122	9	US-10-101-464A-558	Sequence 558, App
32	29	85.3	122	10	US-09-771-383-2	Sequence 2, Appli
33	29	85.3	123	9	US-10-092-154-509	Sequence 509, App
34	29	85.3	123	10	US-09-764-847-509	Sequence 509, App
35	29	85.3	211	9	US-10-050-882-131	Sequence 131, App
36	29	85.3	246	9	US-10-006-856A-338	Sequence 338, App
37	29	85.3	246	9	US-10-006-818A-338	Sequence 338, App
38	29	85.3	246	9	US-10-015-393A-338	Sequence 338, App
39	29	85.3	246	9	US-09-946-374-338	Sequence 338, App
40	29	85.3	246	9	US-10-012-121A-338	Sequence 338, App
41	29	85.3	246	9	US-10-015-869A-338	Sequence 338, App
42	29	85.3	246	9	US-10-006-116A-338	Sequence 338, App
43	29	85.3	246	9	US-10-006-117A-338	Sequence 338, App
44	29	85.3	246	9	US-10-013-913A-338	Sequence 338, App
45	29	85.3	246	9	US-10-017-527A-338	Sequence 338, App
46	29	85.3	246	9	US-10-007-194A-338	Sequence 338, App
47	29	85.3	246	9	US-10-013-430A-338	Sequence 338, App
48	29	85.3	246	9	US-10-011-671A-338	Sequence 338, App
49	29	85.3	246	9	US-10-012-755A-338	Sequence 338, App
50	29	85.3	246	9	US-10-015-386A-338	Sequence 338, App
51	29	85.3	246	9	US-10-011-692A-338	Sequence 338, App
52	29	85.3	285	9	US-09-892-877-134	Sequence 134, App
53	29	85.3	301	9	US-09-948-783-136	Sequence 136, App
54	29	85.3	301	9	US-10-028-072-166	Sequence 166, App
55	29	85.3	301	9	US-10-121-049-166	Sequence 166, App
56	29	85.3	301	9	US-10-123-904-166	Sequence 166, App
57	29	85.3	301	9	US-10-140-470-166	Sequence 166, App
58	29	85.3	301	9	US-10-175-746-166	Sequence 166, App
59	29	85.3	301	9	US-10-176-918-166	Sequence 166, App
60	29	85.3	301	9	US-10-176-921-166	Sequence 166, App
61	29	85.3	301	9	US-10-137-865-166	Sequence 166, App
62	29	85.3	301	9	US-10-142-431-166	Sequence 166, App
63	29	85.3	301	9	US-10-142-431-166	Sequence 166, App
64	29	85.3	301	9	US-10-143-114-166	Sequence 166, App
65	29	85.3	301	9	US-10-140-002-166	Sequence 166, App
66	29	85.3	301	9	US-10-142-419-166	Sequence 166, App
67	29	85.3	301	9	US-10-123-262-166	Sequence 166, App
68	29	85.3	301	9	US-10-142-423-166	Sequence 166, App
69	29	85.3	301	9	US-10-121-050-166	Sequence 166, App
70	29	85.3	301	9	US-10-141-755-166	Sequence 166, App
71	29	85.3	301	9	US-10-143-032-166	Sequence 166, App
72	29	85.3	301	9	US-10-123-108-166	Sequence 166, App
73	29	85.3	301	9	US-10-123-236-166	Sequence 166, App
74	29	85.3	301	9	US-10-123-261-166	Sequence 166, App
75	29	85.3	301	9	US-10-140-921-166	Sequence 166, App

; APPLICANT: Hu, Zengjian
 ; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
 ; FILE REFERENCE: 100086.413C1
 ; CURRENT APPLICATION NUMBER: US/09/769,145
 ; CURRENT FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: US 09/491,078
 ; PRIOR FILING DATE: 2000-01-24
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 18
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
 ; OTHER INFORMATION: peptide with classical cadherin cell adhesion
 ; OTHER INFORMATION: recognition sequence
 ; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
 ; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
 ; OTHER INFORMATION: and/or C-terminal modifications such as amide or
 ; OTHER INFORMATION: ester group
 US-09-769-145-18

Query Match 100.0%; Score 34; DB 9; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.7e+05; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 |||||
 Db 1 CVAHC 5

RESULT 2

US-10-058-821-15
 ; Sequence 15, Application US/10058821
 ; Publication No. US20030087811A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Blaschuk, Orest W.
 ; APPLICANT: Gour, Barbara J.
 ; APPLICANT: Farookhi, Riab
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
 ; FILE REFERENCE: 100086.401C12
 ; CURRENT APPLICATION NUMBER: US/10/058,821
 ; CURRENT FILING DATE: 2002-01-29
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
 ; OTHER INFORMATION: peptide with classical cadherin cell adhesion
 ; OTHER INFORMATION: recognition sequence
 ; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
 ; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
 ; OTHER INFORMATION: and/or C-terminal modifications such as amide or
 ; OTHER INFORMATION: ester group
 US-10-058-821-15

Query Match 100.0%; Score 34; DB 9; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 |||||
 Db 1 CVAHC 5

RESULT 3

US-09-864-761-41470
 ; Sequence 41470, Application US/09864761
 ; Patent No. US20020048763A1

; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
 ; FILE REFERENCE: Aomic-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 41470
 ; LENGTH: 44
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL133295.11
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.8
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.3
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5
 ; OTHER INFORMATION: EST HUMAN HIT: BE245070.1, EVALUATE 5.00e-21
 ; OTHER INFORMATION: SWISSPROT HIT: Q05650, EVALUATE 8.20e+00
 US-09-864-761-41470

Query Match 97.1%; Score 33; DB 10; Length 44;
 Best Local Similarity 80.0%; Pred. No. 40;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 |||||
 Db 30 CVAHC 34

RESULT 4

US-10-046-643-2
; Sequence 2, Application US/10046643
; Patent No. US20020115630A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: 33449 A Human Protease Family Member
; TITLE OF INVENTION: and Uses Therefor
; FILE REFERENCE: MPI2001-016PirCPI(M)
; CURRENT APPLICATION NUMBER: US/10/046,643
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: 60/262,513
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-046-643-2

Query Match 97.1%; Score 33; DB 12; Length 166;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVAHC 5
|:|
Db 21 CIAHC 25

RESULT 5

US-09-776-724A-154
; Sequence 154, Application US/09776724A
; Publication No. US20030050455A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 64 Human Secreted Proteins
; FILE REFERENCE: P2011
; CURRENT APPLICATION NUMBER: US/09/776,724A
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/180,909
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 09/669,688
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/229,982
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: PCT/US98/14613
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/052,661
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,872
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,871
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,874
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,873
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,870
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,875
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/053,440
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,441
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,442
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/056,359
; PRIOR FILING DATE: 1997-08-18

; PRIOR APPLICATION NUMBER: 60/055,725
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,985
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,952
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,989
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,361
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,726
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,724
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,946
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,683
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 154
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-776-724A-154

Query Match 91.2%; Score 31; DB 9; Length 24;

Best Local Similarity 80.0%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVAHC 5
|:|
Db 13 CVSHC 17

RESULT 6

US-09-864-761-34661
; Sequence 34661, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34661
LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AF170801.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
OTHER INFORMATION: SWISSPROT HIT: Q61555, EVALUATE 4.60e+00
OTHER INFORMATION: EST_HUMAN HIT: BF208705.1, EVALUATE 4.00e+00
US-09-864-761-34661

Query Match 91.2%; Score 31; DB 10; Length 47;
Best Local Similarity 80.0%; Pred. No. 90;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|:|
DB 13 CLAHC 17

RESULT 7
US-09-764-877-1632
Sequence 1632, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1632
LENGTH: 68
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (27)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (39)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1632

Query Match 91.2%; Score 31; DB 10; Length 68;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|:|
DB 52 CVSHC 56
RESULT 8
US-09-800-729-179
Sequence 179, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: PZ044PI
CURRENT APPLICATION NUMBER: US/09/800,729
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 179
LENGTH: 96
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-729-179

Query Match 91.2%; Score 31; DB 10; Length 96;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|:|:|
DB 18 CVSHC 22

RESULT 9
US-09-769-145-19
Sequence 19, Application US/09769145
Patent No. US20020168761A1
GENERAL INFORMATION:
APPLICANT: Gour, Barbara J.
APPLICANT: Blaschuk, Orest W.
APPLICANT: Ali, Ammar
APPLICANT: Ni, Feng
APPLICANT: Chen, Zhigang
APPLICANT: Michaud, Stephanie
APPLICANT: Wang, Shoameng
APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: control peptide
OTHER INFORMATION: Cyclic peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-19

Query Match 88.2%; Score 30; DB 9; Length 5;
Best Local Similarity 80.0%; Pred. No. 3.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | |
Db 1 CVGHC 5

RESULT 10

US-09-030-619-218
; Sequence 218, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: McNicol, Patricia J.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION

; FILE REFERENCE: 660081.406

; CURRENT APPLICATION NUMBER: US/09/030.619B

; CURRENT FILING DATE: 1998-02-25

; NUMBER OF SEQ ID NOS: 232

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 218

; LENGTH: 39

; TYPE: PRT

; ORGANISM: Phormia terronovae

US-09-030-619-218

Query Match 88.2%; Score 30; DB 10; Length 39;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | |
Db 16 CAAHC 20

RESULT 11

US-09-030-619-198
; Sequence 198, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: McNicol, Patricia J.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION

; FILE REFERENCE: 660081.406

; CURRENT APPLICATION NUMBER: US/09/030.619B

; CURRENT FILING DATE: 1998-02-25

; NUMBER OF SEQ ID NOS: 232

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 198

; LENGTH: 40

; TYPE: PRT

; ORGANISM: Sacrophaga peregrina

US-09-030-619-198

Query Match 88.2%; Score 30; DB 10; Length 40;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | |
Db 16 CAAHC 20

RESULT 12

US-09-030-619-217
; Sequence 217, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: McNicol, Patricia J.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION

; FILE REFERENCE: 660081.406

; CURRENT APPLICATION NUMBER: US/09/030.619B

; CURRENT FILING DATE: 1998-02-25

; NUMBER OF SEQ ID NOS: 232

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 217

; LENGTH: 40

; TYPE: PRT

; ORGANISM: Phormia terronovae

US-09-030-619-217

Query Match 88.2%; Score 30; DB 10; Length 40;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | |
Db 16 CAAHC 20

RESULT 13

US-09-917-340-73
; Sequence 73, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:

; APPLICANT: Murphy, Christopher J.

; APPLICANT: McNulty, Jonathan F.

; APPLICANT: Reid, Ted W.

; TITLE OF INVENTION: Transplant Media

; FILE REFERENCE: TPLANT-06468

; CURRENT APPLICATION NUMBER: US/09/917.340

; CURRENT FILING DATE: 2001-07-29

; PRIOR APPLICATION NUMBER: 60/221,632

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 60/249,602

; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 60/290,932

; PRIOR FILING DATE: 2001-05-15

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 73

; LENGTH: 40

; TYPE: PRT

; ORGANISM: Aedes aegypti

US-09-917-340-73

Query Match 88.2%; Score 30; DB 10; Length 40;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
| | | |
Db 16 CAAHC 20

RESULT 14

US-09-917-340-94
; Sequence 94, Application US/09917340

; Patent No. US20020090369A1

; GENERAL INFORMATION: Christopher J.

; APPLICANT: Murphy, Christopher J.

; APPLICANT: McAnulty, Jonathan F.

; APPLICANT: Reid, Ted W.

; TITLE OF INVENTION: Transplant Media

; FILE REFERENCE: TPLANT-06468

; CURRENT APPLICATION NUMBER: US/09/917,340

; CURRENT FILING DATE: 2001-07-29

; PRIOR APPLICATION NUMBER: 60/221,632

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 60/249,602

; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 60/290,932

; PRIOR FILING DATE: 2001-05-15

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 94

; LENGTH: 43

; TYPE: PRT

; ORGANISM: Allomyrina dichotoma

; US-09-917-340-94

Query Match 88.2%; Score 30; DB 10; Length 43;

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

Db 20 CAABC 24

RESULT 15

US-10-196-437A-3

; Sequence 3, Application US/10196437A

; Publication No. US20030095983A1

; GENERAL INFORMATION:

; APPLICANT: Webster, Philippa

; APPLICANT: Kelley, James

; TITLE OF INVENTION: USE OF CORTICOTROPH-DERIVED GLYCOPROTEIN

; FILE REFERENCE: 01-16

; CURRENT APPLICATION NUMBER: US/10/196,437A

; CURRENT FILING DATE: 2002-07-15

; PRIOR APPLICATION NUMBER: 60/305,284

; PRIOR FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: Fast-SEQ for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 106

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-196-437A-3

Query Match 88.2%; Score 30; DB 9; Length 106;

Best Local Similarity 80.0%; Pred. No. 2.6e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

Db 34 CVGHC 38

RESULT 16

US-10-196-437A-2

; Sequence 2, Application US/10196437A

; Publication No. US20030095983A1

; GENERAL INFORMATION:

; APPLICANT: Webster, Philippa

; APPLICANT: Kelley, James

; TITLE OF INVENTION: USE OF CORTICOTROPH-DERIVED GLYCOPROTEIN

; FILE REFERENCE: 01-16

; HORMONE TO INDUCE LYPOLYSIS

; HORMONE TO INDUCE LYPOLYSIS

; HORMONE TO INDUCE LYPOLYSIS

; HORMONE TO INDUCE LYPOLYSIS

; HORMONE TO INDUCE LYPOLYSIS

; HORMONE TO INDUCE LYPOLYSIS

; CURRENT APPLICATION NUMBER: US/10/196,437A

; CURRENT FILING DATE: 2002-07-15

; PRIOR APPLICATION NUMBER: 60/305,284

; PRIOR FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: Fast-SEQ for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 129

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-196-437A-2

Query Match 88.2%; Score 30; DB 9; Length 129;

Best Local Similarity 80.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

Db 57 CVGHC 61

RESULT 17

US-09-970-711-34

; Sequence 34, Application US/09970711

; Patent No. US20020081279A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Adam

; APPLICANT: Cotten, Matthew

; APPLICANT: Chioocca, Susanna

; APPLICANT: Kurzbauer, Robert

; APPLICANT: Schaffner, Gotthold

; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus

; FILE REFERENCE: 0652.1800001

; CURRENT APPLICATION NUMBER: US/09/970,711

; CURRENT FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 09/171,461

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: PCT/EP97/01944

; PRIOR FILING DATE: 1997-04-18

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 34

; LENGTH: 210

; TYPE: PRT

; ORGANISM: CELO VIRUS

; FEATURE:

; OTHER INFORMATION: Position: 4462..5094/note=ORF12

; US-09-970-711-34

Query Match 88.2%; Score 30; DB 10; Length 210;

Best Local Similarity 80.0%; Pred. No. 4.6e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

Db 86 CTAHC 90

RESULT 18

US-09-815-242-10567

; Sequence 10567, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Hasebeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 1410
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 10567
 LENGTH: 906
 TYPE: PRT
 ORGANISM: Enterococcus faecalis
 US-09-815-242-10567

Query Match 88.2%; Score 30; DB 10; Length 906;
 Best Local Similarity 80.0%; Pred. No. 1.6e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
 DB 72 CTAHC 76

RESULT 19

US-10-172-620-16
 Sequence 16, Application US/10172620
 Publication No. US20030053995A1
 GENERAL INFORMATION:
 APPLICANT: Hung, Mien-Chie
 APPLICANT: Lin, Shiao-Yih
 TITLE OF INVENTION: Methods and Compositions for Inhibiting EGF Receptor
 FILE REFERENCE: UTSC:720US
 CURRENT APPLICATION NUMBER: US/10/172,620
 CURRENT FILING DATE: 2002-06-14
 PRIOR APPLICATION NUMBER: US:60/298,579
 PRIOR FILING DATE: 2001-06-15
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 16
 LENGTH: 1342
 TYPE: PRT
 ORGANISM: Human
 US-10-172-620-16

Query Match 88.2%; Score 30; DB 9; Length 1342;
 Best Local Similarity 80.0%; Pred. No. 2.2e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
 DB 529 CVTHC 533

RESULT 20

US-09-764-891-4427
 Sequence 4427, Application US/09764891
 Publication No. US20030077808A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PC006
 CURRENT APPLICATION NUMBER: US/09/764,891

CURRENT FILING DATE: 2001-01-17
 Prior application data removed - consult PALM or file wrapper
 NUMBER OF SEQ ID NOS: 10231
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 4427
 LENGTH: 42
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-764-891-4427

Query Match 85.3%; Score 29; DB 9; Length 42;
 Best Local Similarity 80.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
 DB 22 CVLHC 26

RESULT 21

US-10-178-213-378
 Sequence 378, Application US/10178213
 Publication No. US20030041348A1
 GENERAL INFORMATION:
 APPLICANT: Simmons, Carl R.
 APPLICANT: Navarro Acevedo, Pedro A.
 APPLICANT: Harvell, Leslie
 APPLICANT: Cahoon, Rebecca
 APPLICANT: McCutchen, Billy Fred
 APPLICANT: Lu, Albert
 APPLICANT: Herrmann, Rafael
 APPLICANT: Wong, James
 TITLE OF INVENTION: Defensin Polynucleotides and Methods of
 TITLE OF INVENTION: Use
 FILE REFERENCE: 35718/246703
 CURRENT APPLICATION NUMBER: US/10/178,213
 CURRENT FILING DATE: 2002-06-21
 PRIOR APPLICATION NUMBER: 60/300,152
 PRIOR FILING DATE: 2001-06-22
 PRIOR APPLICATION NUMBER: 60/300,241
 PRIOR FILING DATE: 2001-06-22
 NUMBER OF SEQ ID NOS: 469
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 378
 LENGTH: 47
 TYPE: PRT
 ORGANISM: Parthenium argentatum
 US-10-178-213-378

Query Match 85.3%; Score 29; DB 9; Length 47;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
 DB 43 CVKHC 47

RESULT 22

US-10-178-213-288
 Sequence 288, Application US/10178213
 Publication No. US20030041348A1
 GENERAL INFORMATION:
 APPLICANT: Simmons, Carl R.
 APPLICANT: Navarro Acevedo, Pedro A.
 APPLICANT: Harvell, Leslie
 APPLICANT: Cahoon, Rebecca
 APPLICANT: McCutchen, Billy Fred
 APPLICANT: Lu, Albert
 APPLICANT: Herrmann, Rafael
 APPLICANT: Wong, James
 TITLE OF INVENTION: Defensin Polynucleotides and Methods of
 TITLE OF INVENTION: Use

```
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 288
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-178-213-288

Query Match      85.3%; Score 29; DB 9; Length 52;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CVAHC 5
      |||||
Db      20 CVEHC 24

RESULT 23
; Sequence 445, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 445
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-391-445

Query Match      85.3%; Score 29; DB 9; Length 63;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CVAHC 5
      |||||
Db      40 CVLHC 44

RESULT 24
US-10-091-572-409
; Sequence 409, Application US/10091572
; Publication No. US20030054373A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA118C1
; CURRENT APPLICATION NUMBER: US/10/091,572
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,850
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
```

; PRIOR APPLICATION NUMBER: 60/236,367
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: 60/237,039
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/237,038
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/236,370
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: 60/236,802
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/237,037
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/237,040
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/240,960
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/239,935
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: 60/239,937
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: 60/241,787
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/246,474
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/246,532
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/249,216
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,210
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/226,681
 ; PRIOR FILING DATE: 2000-08-22
 ; PRIOR APPLICATION NUMBER: 60/225,759
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/225,213
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/227,182
 ; PRIOR FILING DATE: 2000-08-22
 ; PRIOR APPLICATION NUMBER: 60/225,214
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/230,438
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/215,135
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: 60/225,266
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/249,218
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,208
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,213
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,212
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,207
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,245
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,244
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,217
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,211
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,215
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,264
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,214
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,297

; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/232,400
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/231,242
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/232,081
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/232,080
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/231,414
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/231,244
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/233,064
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/233,063
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,397
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,399
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,401
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/241,808
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,826
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,786
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,221
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/246,475
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/231,243
 ; PRIOR FILING DATE: 2000-09-08

Query Match 85.3%; Score 29; DB 9; Length 73;
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
 DB 5 CVAHC 9

RESULT 25
 US-10-106-698-6006
 ; Sequence 6006, Application US/10106698
 ; Publication No. US20030109690A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
 ; FILE REFERENCE: PA005P1
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: US/10/106,698
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/157,137
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: US 60/163,280
 ; PRIOR FILING DATE: 1999-11-03
 ; NUMBER OF SEQ ID NOS: 8564
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 6006
 ; LENGTH: 75
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-106-698-6006

Query Match 85.3%; Score 29; DB 9; Length 75;
 Best Local Similarity 60.0%; Pred. No. 2.8e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVARC 5
|: ||
Db 17 CIVRC 21

Search completed: June 20, 2003, 21:02:37
Job time : 14.8409 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 12.7273 Seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-18
Perfect score: 34
Sequence: 1 CVARC 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 283224
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	67	2 E81135	hypothetical prote
2	34	100.0	305	2 G70382	phosphotidylglycer
3	34	100.0	438	2 H82734	acetylglutamate ki
4	34	100.0	547	2 T51761	cyclohex-1-ene-1-c
5	31	91.2	88	2 T34429	hypothetical prote
6	31	91.2	173	2 T32912	hypothetical prote
7	31	91.2	182	2 B70315	hypothetical prote
8	31	91.2	340	2 T35406	hypothetical prote
9	31	91.2	384	2 C64752	yaga protein - Bac
10	31	91.2	400	2 S76066	hypothetical prote
11	31	91.2	405	2 T29639	hypothetical prote
12	31	91.2	451	2 B96774	hypothetical prote
13	31	91.2	844	2 T37690	hypothetical prote
14	30	88.2	43	2 J32554	holotricin 1 - Hol
15	30	88.2	55	1 NTRNB4	neurotoxin B-IV -
16	30	88.2	55	1 NTRNB2	neurotoxin B-II -
17	30	88.2	84	2 JX0332	tensacin 1 precursor
18	30	88.2	94	2 A31792	sapsacin precursor
19	30	88.2	94	2 S12558	defensin A precursor
20	30	88.2	159	2 A64338	ferripyochelin bin
21	30	88.2	292	2 D91096	probable dehydroge
22	30	88.2	292	2 H85941	probable dehydroge
23	30	88.2	292	2 C65070	hypothetical prote
24	30	88.2	323	2 D82345	UDP-glucose 4-epim
25	30	88.2	327	2 A44013	gonadotropin-relea
26	30	88.2	327	2 I60169	gonadotropin-relea
27	30	88.2	328	2 JN0882	gonadotropin-relea
28	30	88.2	328	2 JCI353	gonadotropin-relea
29	30	88.2	357	2 F82982	catabolic alanine

30	30	88.2	374	2	F81693
31	30	88.2	465	2	H86482
32	30	88.2	468	2	T50873
33	30	88.2	520	2	S14599
34	30	88.2	520	2	S14600
35	30	88.2	520	2	S14598
36	30	88.2	522	2	B85165
37	30	88.2	534	2	S60205
38	30	88.2	544	2	S41626
39	30	88.2	550	1	VGIHD6
40	30	88.2	555	2	T21028
41	30	88.2	620	2	G71413
42	30	88.2	676	2	T22532
43	30	88.2	748	2	S66129
44	30	88.2	899	2	G02428
45	30	88.2	910	2	S68983
46	30	88.2	915	2	JC6148
47	30	88.2	1154	1	VGIHIB
48	30	88.2	1162	2	S07421
49	30	88.2	1162	2	S14939
50	30	88.2	1162	2	S14940
51	30	88.2	1299	2	T43251
52	30	88.2	1305	2	T31096
53	30	88.2	1339	2	JC4387
54	30	88.2	1342	2	A36223
55	30	88.2	4544	1	S02392
56	29	85.3	23	2	B38671
57	29	85.3	65	2	F82580
58	29	85.3	77	2	A82463
59	29	85.3	93	2	D82474
60	29	85.3	104	2	G49897
61	29	85.3	111	2	C72514
62	29	85.3	122	2	H70370
63	29	85.3	131	2	F97868
64	29	85.3	141	2	S30832
65	29	85.3	161	2	F86191
66	29	85.3	182	2	H84939
67	29	85.3	190	2	G69764
68	29	85.3	193	2	T17449
69	29	85.3	193	2	AI0231
70	29	85.3	208	2	T27749
71	29	85.3	208	2	T27748
72	29	85.3	244	2	D70851
73	29	85.3	361	2	T29571
74	29	85.3	374	2	D71538
75	29	85.3	399	2	S68805

ALIGNMENTS

RESULT 1

E81135

hypothetical protein NMB0988 [imported] - Neisseria meningitidis (strain MC58 serog.
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: E81135
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
ri, H.; Qin, H.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty,
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: E81135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <FET>
A:Cross-references: GB:AE002449; GB:AE002098; NID:G7226218; PIDN:AAF41391.1; PID:G72
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0988

iron-sulfur cofact
protein FSJ5.11 [1
hypothetical prote
E2 glycoprotein -
E2 glycoprotein pr
E2 glycoprotein -
hypothetical prote
phosphonoacetaldeh
epike protein chai
E2 glycoprotein pr
hypothetical prote
hypothetical 7K pr
hypothetical prote
subintegrin (EC 3.
subtilisin-like pr
subtilisin - bovine
subtilisin-like pr
E2 glycoprotein pr
E2 glycoprotein pr
E2 glycoprotein pr
E2 glycoprotein pr
furin (EC 3.4.21.7
cyclin G-associate
epidermal growth f
kinase-related tra
alpha-2-macroglobu
peptidylglycine mo
hypothetical prote
conserved hypothet
hypothetical prote
conserved hypothet
hypothetical prote
holo-[acyl-carrier
hypothetical prote
hypothetical prote
inorganic diphosph
transcription regu
hypothetical 21.9K
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
iron-sulfur cofact
pectin acetyllester

Query Match 100.0%; Score 34; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
 |||||
 DB 9 CVAHC 13

RESULT 2

G70382

phosphotidylglycerophosphate synthase - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C:Accession: G70382

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oviatt, D.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: G70382

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-305 <AQF>

A:Cross-references: GB:AE000715; NID:g2983460; PIDN:AAC07048.1; PID:g2983468; GB:AE00065

A:Experimental source: strain VF5

C:Genetics:

A:Gene: pgsA2

Query Match 100.0%; Score 34; DB 2; Length 305;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
 |||||
 DB 296 CVAHC 300

RESULT 3

H82734

acetylglutamate kinase XF1001 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: H82734

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: H82734

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-438 <STM>

A:Cross-references: GB:AE003938; GB:AE003849; NID:g9105935; PIDN:AAF83811.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Pacincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva

M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1001

Query Match 100.0%; Score 34; DB 2; Length 438;

Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
 |||||
 DB 425 CVAHC 429

RESULT 4

T51761

cyclohex-1-ene-1-carboxylate CoA ligase [imported] - Rhodopseudomonas palustris

C:Species: Rhodopseudomonas palustris

C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 03-Nov-2000

C:Accession: T51761

R:England, P.G.; Pelletier, D.A.; Dispensa, M.; Gibson, J.; Harwood, C.S.

Proc. Natl. Acad. Sci. U.S.A. 94, 6484-6489, 1997

A:Title: A cluster of bacterial genes for anaerobic benzene ring biodegradation.

A:Reference number: Z54449; MUID:97322399; PMID:9177244

A:Accession: T51761

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-547 <EGL>

A:Cross-references: EMBL:U75363; PIDN:AAC23919.1

A:Experimental source: strain CGA009

C:Genetics:

A:Gene: aliA

C:Function:

A:Description: converts cyclohexanecarboxylate to cyclohexanecarboxyl-CoA

C:Superfamily: acetate-CoA ligase homology

F:76-538/Domain: acetate-CoA ligase homology <ACL>

Query Match 100.0%; Score 34; DB 2; Length 547;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
 |||||
 DB 33 CVAHC 37

RESULT 5

T34429

hypothetical protein F55C7.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34429

R:Du, Z.; Le, T.

submitted to the EMBL Data Library, November 1996

A:Description: The sequence of C. elegans eosmid F55C7.

A:Reference number: Z21524

A:Accession: T34429

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-88 <DUZ>

A:Cross-references: EMBL:U80436; PIDN:AAC71107.1; GSPDB:GN000019; CESP:F55C7.2

A:Experimental source: strain Bristol N2; clone F55C7

C:Genetics:

A:Gene: CESP:F55C7.2

A:Map position: 1

A:Introns: 18/3

Query Match 91.2%; Score 31; DB 2; Length 88;
 Best Local Similarity 80.0%; Pred. No. 55;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
 |||||
 DB 72 CVAHC 76

RESULT 6

T32912

hypothetical protein C54G6.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T32912
R;Pin-Wollam, A.; Graves, T.; Ozersky, P.
submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid C54G6.
A;Reference number: Z21245
A;Accession: T32912
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-173 <TIN>
A;Cross-references: EMBL:AF043698; PIDN:AAB97561.1; GSPDB:GN00019; CESP:C54G6.3
A;Experimental source: strain Bristol N2; clone C54G6
C;Genetics:
A;Gene: CESP:C54G6.3
A;Map position: 1
A;Introns: 17/21; 51/3; 94/1; 169/2

Query Match 91.2%; Score 31; DB 2; Length 173;
Best Local Similarity 80.0%; Pred. No. 92;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|||
Db 101 CVSHC 105

RESULT 7
B70315
hypothetical protein aq_159 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
A;Accession: B70315
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O.V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: B70315
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-182 <AQF>
A;Cross-references: GB:AE000676; NID:g2982884; PIDN:AAC06513.1; PID:g2982894; GB:AE00065
A;Experimental source: strain VFS
C;Genetics:
A;Gene: aq_159

Query Match 91.2%; Score 31; DB 2; Length 182;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|||
Db 98 CVSHC 102

RESULT 8
T35406
hypothetical protein SC6A5.15 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
A;Accession: T35406
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21577
A;Accession: T35406
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-340 <OLI>
A;Cross-references: EMBL:AL049485; PIDN:CAB39699.1; GSPDB:GN00070; SCOEDB:SC6A5.15
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC6A5.15

Query Match 91.2%; Score 31; DB 2; Length 340;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|||
Db 306 CVAHC 310

RESULT 9
C64752
yagA protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
A;Accession: C64752
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A84720; MUID:97426617; PMID:9278503
A;Accession: C64752
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-384 <BLAT>
A;Cross-references: GB:AE000134; GB:U00096; NID:g1786454; PIDN:AAC73370.1; PID:g178
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yagA

Query Match 91.2%; Score 31; DB 2; Length 384;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|||
Db 175 CLAHC 179

RESULT 10
S76066
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
A;Accession: S76066
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Y.
DNA Res 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76066
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-400 <XAN>
A;Cross-references: EMBL:D63999; GB:AB001339; NID:gi001396; PIDN:BAAL0044.1; PID:gi0
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 91.2%; Score 31; DB 2; Length 400;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
|||
Db 248 CVSHC 252

RESULT 11
T29639
hypothetical protein F49B8.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C;Accession: T29639
R;Murray, J.; Woldmann, P.
submitted to the EMBL Data Library, June 1996
A;Description: The sequence of *C. elegans* cosmid F49E8.
A;Reference number: Z20657

A;Accession: T29639
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-405 <MUR>

A;Cross-references: EMBL:U61949; PIDN:AAB03150.1; GSPDB:GN00022; CESP:F49E8.2
A;Experimental source: strain Bristol N2, clone F49E8

C;Genetics:

A;Gene: CESP:F49E8.2

A;Map position: 4

A;Introns: 71/3; 135/3; 172/1; 207/3; 284/2; 347/3

C;Superfamily: Caenorhabditis elegans hypothetical protein F49E8.2

Query Match 91.2%; Score 31; DB 2; Length 405;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5

DB 6 CVSHC 10

RESULT 12

B96774

hypothetical protein F1M20.19 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: B96774

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansan, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B96774

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-451 <STO>

A;Cross-references: GB:AE005173; NID:96539253; PIDN:AAF15923.1; GSPDB:GN00141

C;Genetics:

A;Gene: F1M20.19

A;Map position: 1

Query Match 91.2%; Score 31; DB 2; Length 451;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5

DB 107 CLAHC 111

RESULT 13

T37690

hypothetical protein SPAC14C.05c - fission yeast (*Schizosaccharomyces pombe*)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T37690

R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1997

A;Reference number: Z21737

A;Accession: T37690

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-844 <DEV>

A;Cross-references: EMBL:Z98596; PIDN:CAB11198.1; GSPDB:GN00066; SPDB:SPAC14C.05c

A;Experimental source: strain 972h-, cosmid c14C4

C;Genetics:

A;Gene: SPDB:SPAC14C.05c

A;Map position: 1

Query Match 91.2%; Score 31; DB 2; Length 844;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5

DB 745 CVSHC 749

RESULT 14

JC2554

holotricin 1 - Holotrichia diomphalia

C;Species: Holotrichia diomphalia

C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 17-Mar-1999

C;Accession: JC2554

R;Lee, S.Y.; Moon, H.J.; Kawabata, S.; Kurata, S.; Natori, S.; Lee, B.L.

Biol. Pharm. Bull. 18, 457-459, 1995

A;Title: A sapecin homologue of Holotrichia diomphalia: Purification, sequencing and

A;Reference number: JC2554; MUID:96031330; PMID:7550103

A;Accession: JC2554

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-43 <LEE>

C;Superfamily: sapecin precursor

Query Match 88.2%; Score 30; DB 2; Length 43;

Best Local Similarity 80.0%; Pred. No. 48;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5

DB 20 CAHC 24

RESULT 15

NTN84

neurotoxin B-IV - ribbon worm (*Cerebratulus lacteus*)

C;Species: *Cerebratulus lacteus* (milky ribbon-worm)

C;Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 04-Oct-1996

C;Accession: A92340; A01788

R;Blumenthal, K.M.; Keim, P.S.; Heinrichson, R.L.; Ken, W.R.

J. Biol. Chem. 256, 9063-9067, 1981

A;Title: Structure and action of heteronemertine polypeptide toxins. Amino acid seq

A;Reference number: A92340; MUID:81264350; PMID:7263698

A;Accession: A92340

A;Molecule type: protein

A;Residues: 1-55 <BLU>

R;Blumenthal, K.M.; Ken, W.R.

J. Biol. Chem. 251, 6025-6029, 1976

A;Title: Structure and action of heteronemertine polypeptide toxins. Primary struct

A;Reference number: A92195; MUID:77006160; PMID:972152

C;Superfamily: neurotoxin B-IV

C;Keywords: hydroxyproline; neurotoxin

F;10/Modified site: hydroxyproline (pro) #status experimental

F;12-48,16-52,23-41,26-37/Diulfide bonds: #status experimental

Query Match 88.2%; Score 30; DB 1; Length 55;

Best Local Similarity 80.0%; Pred. No. 58;

Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CVAHC 5

DB 37 CAHC 41

RESULT 16

NTN82
neurotoxin B-II - ribbon worm (Cerebratulus lacteus)
C:Species: Cerebratulus lacteus (milky ribbon-worm)
C:Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 04-Oct-1996
A:Accession: A01789
R:Blumenthal, K.M.; Keim, P.S.; Heinrikson, R.L.; Ken, W.R.
J. Biol. Chem. 256, 9063-9067, 1981
A:Title: Structure and action of heteronemertine polypeptide toxins. Amino acid sequence
A:Reference number: A92340; MUID:81264350; PMID:7263698
A:Accession: A01789
A:Molecule type: protein
A:Residues: 1-55 <BLU>
C:Superfamily: neurotoxin B-IV
C:Keywords: hydroxyproline; neurotoxin
F:10/Modified site: hydroxyproline (Pro) #status experimental
F:12-48,16-52,23-41,26-37/Disulfide bonds: #status experimental

Query Match 88.2%; Score 30; DB 1; Length 55;
Best Local Similarity 80.0%; Pred. No. 58;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
|
|
|
|
Db 37 CAABC 41

RESULT 17

JX0332
teneurin 1 precursor - yellow mealworm
C:Species: Tenebrio molitor (yellow mealworm)
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 20-Jun-2000
A:Accession: JX0332
R:Moore, H.J.; Lee, S.Y.; Kurata, S.; Natori, S.; Lee, B.L.
J. Biochem. 116, 53-58, 1994
A:Title: Purification and molecular cloning of cDNA for an inducible antibacterial protease
A:Reference number: JX0332; MUID:95096025; PMID:7798186
A:Accession: JX0332
A:Molecule type: mRNA
A:Residues: 1-84 <MOO>
A:Cross-references: GB:D17670; NID:g1235940; PIDN:BA04552.1; PID:g1235941
A:Experimental source: larva
C:Comment: Teneurin 1 is an antibacterial protein induced in the hemolymph of larvae of B
C:Superfamily: sapecin precursor
F:1-41/Domain: signal sequence #status predicted <SIG>
F:42-84/Product: teneurin 1 #status predicted <MAT>
F:44-75,61-81,65-83/Disulfide bonds: #status predicted

Query Match 88.2%; Score 30; DB 2; Length 84;
Best Local Similarity 80.0%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
|
|
|
|
Db 61 CAABC 65

RESULT 18

A31792
sapecin precursor - flesh fly (Sarcophaga peregrina)
N:Alternate names: antibacterial protein
C:Species: Sarcophaga peregrina
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
A:Accession: A31792; A31791
R:Matsumura, K.; Natori, S.
J. Biol. Chem. 263, 17117-17121, 1988
A:Title: Molecular cloning of cDNA for sapecin and unique expression of the sapecin gene
A:Reference number: A31792; MUID:89034216; PMID:3182837
A:Accession: A31792
A:Molecule type: mRNA
A:Residues: 1-94 <MAT>
A:Cross-references: GB:J04053; NID:g161266; PIDN:AAA29984.1; PID:g161267
A:Experimental source: K.; Natori, S.

J. Biol. Chem. 263, 17112-17116, 1988
A:Title: Purification of three antibacterial proteins from the culture medium of NIH
A:Reference number: A92687; MUID:89034215; PMID:3182836
A:Accession: A31791
A:Molecule type: protein
A:Residues: 55-94 <MA2>
C:Superfamily: sapecin precursor

Query Match 88.2%; Score 30; DB 2; Length 94;
Best Local Similarity 80.0%; Pred. No. 87;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
|
|
|
|
Db 70 CAABC 74

RESULT 19

S12558
defensin A precursor - nestling-sucking blowfly
N:Alternate names: phormicin precursor
C:Species: Protophormia terraenovae (nestling-sucking blowfly)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 15-Oct-1999
A:Accession: S12558; A32219; B32219
R:Dimarcq, J.L.; Zachary, D.; Hoffmann, J.A.; Hoffmann, D.; Reichhart, J.M.
EMBO J. 9, 2507-2515, 1990
A:Title: Insect immunity: expression of the two major inducible antibacterial peptides
A:Reference number: S12558; MUID:90316108; PMID:2369900
A:Accession: S12558
A:Molecule type: mRNA
A:Residues: 1-94 <DIM>
A:Cross-references: EMBL:X55546; NID:g10923; PIDN:CAA39152.1; PID:g10924
R:Lambert, J.; Keppe, E.; Dimarcq, J.L.; Wicker, C.; Reichhart, J.M.; Dunbar, B.; Le
Proc. Natl. Acad. Sci. U.S.A. 86, 262-266, 1989
A:Title: Insect immunity: isolation from immune blood of the dipteran Phormia terraenovae
A:Reference number: A32219; MUID:89098894; PMID:2911573
A:Accession: A32219
A:Molecule type: protein
A:Residues: 55-94 <LAM>
A:Note: peptide A
A:Accession: B32219
A:Molecule type: preliminary
A:Status: preliminary
A:Residues: 55-85, 'R', 87-94 <LA2>
A:Note: peptide B
C:Superfamily: sapecin precursor
C:Keywords: antibacterial
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-54/Domain: propeptide #status predicted <PRO>
F:55-94/Product: defensin A #status experimental <MAT>

Query Match 88.2%; Score 30; DB 2; Length 94;
Best Local Similarity 80.0%; Pred. No. 87;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
|
|
|
|
Db 70 CAABC 74

RESULT 20

A64338
ferripyochelin binding protein homolog - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
A:Accession: A64338
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: A64338
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-159 <BUL>
 A:Cross-references: GB:U67485; GB:L77117; NID:G1591020; PIDN:AA988291.1; PID:G1591027; T
 C:Superfamily: ferricytochrome binding protein
 A:Map position: REV286504-286025
 C:Genetics:
 A:Map position: REV286504-286025
 C:Superfamily: ferricytochrome binding protein
 Query Match 88.2%; Score 30; DB 2; Length 159;
 Best Local Similarity 80.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 DB 51 CVAHC 55
 RESULT 21
 D91096
 probable dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 050
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 24-Aug-2001
 C:Accession: D91096
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: D91096
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA837163.1; PID:G13363212; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC83740
 C:Superfamily: carbon monoxide dehydrogenase medium chain
 Query Match 88.2%; Score 30; DB 2; Length 292;
 Best Local Similarity 80.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 DB 199 CVAHC 203
 RESULT 22
 H85941
 probable dehydrogenase ygeT [imported] - Escherichia coli (strain O157:H7, substrain EDL
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85941
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85941
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <STO>
 A:Cross-references: GB:AE005174; NID:G12517385; PIDN:AA857996.1; GSPDB:GN00145; UWGP:Z42
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: ygeT
 C:Superfamily: carbon monoxide dehydrogenase medium chain
 Query Match 88.2%; Score 30; DB 2; Length 292;
 Best Local Similarity 80.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
 DB 199 CVAHC 203
 RESULT 23
 C65070
 hypothetical protein b2867 - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: C65070
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: C65070
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-292 <BLAT>
 A:Cross-references: GB:AE000370; GB:U00096; NID:G2367170; PIDN:AA75905.1; PID:G178
 A:Experimental source: strain K-12, substrain MG1655
 C:Superfamily: carbon monoxide dehydrogenase medium chain
 Query Match 88.2%; Score 30; DB 2; Length 292;
 Best Local Similarity 80.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 DB 199 CVAHC 203
 RESULT 24
 D82345
 UDP-glucose 4-epimerase VC0262 [imported] - Vibrio cholerae (strain N16961 serogroup
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: D82345
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson,
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sella
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: D82345
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <HEI>
 A:Cross-references: GB:AE004114; GB:AE003852; NID:G9654662; PIDN:AAF93437.1; GSPDB:G
 A:Experimental source: serogroup O1, strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0262
 A:Map position: 1
 Query Match 88.2%; Score 30; DB 2; Length 323;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVAHC 5
 DB 69 CVAHC 73
 RESULT 25
 A44013
 gonadotropin-releasing hormone receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000
 C:Accession: A44013; A45101
 R:Autsumi, M.; Zhou, W.; Millar, R.P.; Mellon, P.L.; Roberts, J.L.; Flanagan, C.A.
 Mol. Endocrinol. 6, 1163-1169, 1992
 A:Title: Cloning and functional expression of a mouse gonadotropin-releasing hormon

A:Reference number: A44013; MUID:92375111; PMID:1324422
A:Accession: A44013
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-327 <TSU>
A:Experimental source: gonadotrope cell line alpha T3-1
A>Note: sequence extracted from NCBI backbone (NCBIN:111838, NCBIP:111839)
R:Reinhart, J.; Mertz, L.M.; Catt, K.J.
J. Biol. Chem. 267, 21281-21284, 1992
A:Title: Molecular cloning and expression of cDNA encoding the murine gonadotropin-releasing hormone receptor type 1
A:Reference number: A45101; MUID:93016057; PMID:1328228
A:Accession: A45101
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-249, 'L', 251-327 <REI>
A:Experimental source: alpha T3 gonadotroph cell line
A>Note: sequence extracted from NCBI backbone (NCBIP:116690)
C:Superfamily: vertebrate rhodopsin

Query Match 88.2%; Score 30; DB 2; Length 327;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
|||
Db 195 CVTHC 199

Search completed: June 20, 2003, 20:59:06
Job time : 14.7273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 ; Search time 6.02273 Seconds
(without alignments)
34.433 Million cell updates/sec

Title: US-10-105-008-18
Perfect score: 34
Sequence: 1 CVAHC 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	91.2	384	1 YAGA_ECOLI	P37007 escherichia
2	31	91.2	844	1 YD25 SCHPO	O13712 schizosacch
3	31	91.2	1069	1 C110 HUMAN	Q9nzb2 homo sapien
4	30	88.2	40	1 DEPB_AEDAE	P81602 aedes aegyp
5	30	88.2	43	1 DEPI_ALLDI	Q10745 allomyrina
6	30	88.2	55	1 NXB2_CERLA	P01526 cerebratulu
7	30	88.2	55	1 NXB4_CERLA	P01525 cerebratulu
8	30	88.2	79	1 DEFI_STOCA	O16136 stomoxys ca
9	30	88.2	84	1 DEFI_TENMO	Q27023 tenebrio mo
10	30	88.2	94	1 DEFI_PROTE	P10891 protophormi
11	30	88.2	94	1 SARP_SARPE	P18313 sarcophaga
12	30	88.2	97	1 DEFI_STOCA	O16137 stomoxys ca
13	30	88.2	99	1 DEFC_AEDAE	P81603 aedes aegyp
14	30	88.2	102	1 DEFI_ANOGA	Q17027 anopheles g
15	30	88.2	159	1 Y304 METJA	Q57752 methanococ
16	30	88.2	292	1 XDHE_ECO57	Q8x6c5 escherichia
17	30	88.2	292	1 XDHE_ECOLI	Q46800 escherichia
18	30	88.2	327	1 GRHR_MOUSE	Q01776 mus musculu
19	30	88.2	327	1 GRHR_RAT	P30969 rattus norv
20	30	88.2	328	1 GRHR_BOVIN	P32236 bos taurus
21	30	88.2	328	1 GRHR_HORSE	O18821 equus cabal
22	30	88.2	328	1 GRHR_HUMAN	P30968 homo sapien
23	30	88.2	328	1 GRHR_SHEEP	P32237 ovis aries
24	30	88.2	348	1 O3A4 HUMAN	P47883 pseudomonas
25	30	88.2	357	1 ALR2_PSEAE	Q9htq2 pseudomonas
26	30	88.2	520	1 VGL2_IBVU1	P30206 avian infec
27	30	88.2	520	1 VGL2_IBVU2	P30207 avian infec
28	30	88.2	520	1 VGL2_IBVU3	P30208 avian infec
29	30	88.2	550	1 VGL2_IBVD3	P17662 avian infec
30	30	88.2	910	1 AUXI_BOVIN	Q27974 bos taurus
31	30	88.2	913	1 PKC5 HUMAN	Q92824 homo sapien
32	30	88.2	915	1 LIO_DROME	Q24118 drosophila
33	30	88.2	1154	1 VGL2_IBVD2	P12722 avian infec

34	30	88.2	1162	1 VGL2_IBVB	P11223 avian infec
35	30	88.2	1162	1 VGL2_IBVM	P12651 avian infec
36	30	88.2	1163	1 VGL2_IBV6	P05135 avian infec
37	30	88.2	1305	1 GAK_RAT	P97874 rattus norv
38	30	88.2	1311	1 GAK_HUMAN	O14976 homo sapien
39	30	88.2	1339	1 ERB3_RAT	Q62799 rattus norv
40	30	88.2	1342	1 ERB3_HUMAN	P21860 homo sapien
41	30	88.2	4544	1 LRP1_BUCAP	Q07954 homo sapien
42	29	85.3	92	1 Y274_BUCAP	P42398 buchnera ap
43	29	85.3	122	1 ACP5_AQUAE	O66995 aquifex aeo
44	29	85.3	141	1 YEF5_YEAST	P32616 saccharomyc
45	29	85.3	182	1 IPYR_BUCAI	P57190 buchnera ap
46	29	85.3	208	1 YS23_CAEEL	Q09364 caenorhabdi
47	29	85.3	208	1 YS23_CAEEL	Q09365 caenorhabdi
48	29	85.3	455	1 YKYL_CAEEL	O19910 caenorhabdi
49	29	85.3	656	1 SL54_MOUSE	Q9et37 mus musculu
50	29	85.3	659	1 SL54_MOUSE	Q9ny91 homo sapien
51	29	85.3	660	1 SL54_PIG	P31636 sus scrofa
52	29	85.3	749	1 NIBL_MOUSE	Q8rlf1 mus musculu
53	29	85.3	793	1 FURI_MOUSE	P23188 mus musculu
54	29	85.3	794	1 FURI_HUMAN	P09958 homo sapien
55	29	85.3	797	1 FURI_BOVIN	Q28193 bos taurus
56	29	85.3	1062	1 NAL2_HUMAN	Q9nx02 homo sapien
57	29	85.3	1323	1 LT23_CAEEL	P24348 caenorhabdi
58	29	85.3	2146	1 INSR_DROME	P09208 drosophila
59	28	82.4	20	1 SC18_MESMA	P58570 mesobuthus
60	28	82.4	33	1 Y645_BUTOC	P04099 buthus occi
61	28	82.4	58	1 Y645_TREPA	O83651 treponema p
62	28	82.4	64	1 SCX4_ANDAU	P45658 androctonus
63	28	82.4	64	1 SCX4_MESMA	P58328 mesobuthus
64	28	82.4	66	1 SCX1_MESMA	P15224 orthochirus
65	28	82.4	66	1 SCX3_BUTOM	P13488 buthus occi
66	28	82.4	67	1 SC13_LEIOH	P56678 leirus qui
67	28	82.4	67	1 SCX3_ORTSC	P15225 orthochirus
68	28	82.4	74	1 SC17_MESMA	Q9njc4 mesobuthus
69	28	82.4	79	1 SCX9_MESMA	P45698 mesobuthus
70	28	82.4	83	1 SCX1_ANDAU	P01479 androctonus
71	28	82.4	84	1 SC13_MESMA	Q9njc8 mesobuthus
72	28	82.4	84	1 SCX3_ANDAU	P01480 androctonus
73	28	82.4	84	1 SCXB_MESMA	Q9n682 mesobuthus
74	28	82.4	85	1 SC12_MESMA	Q99q43 mesobuthus
75	28	82.4	193	1 LPCA_BUCAI	P57338 buchnera ap

ALIGNMENTS

RESULT 1
YAGA_ECOLI
ID YAGA_ECOLI STANDARD; PRT; 384 AA.
AC P37007; P77312; 384 AA.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yaga.
GN YAGA OR B0267
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA MEDLINE=97426617; PubMed=9278503;
RX Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.,
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;

RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.:
RT "Systematic sequencing of the *Escherichia coli* genome: analysis of the
RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RP Davis K., Pederspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.:
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE OF 245-384 FROM N.A.
RP MEDLINE=94156832; PubMed=8113168;
RX Volkert M.R., Loewen P.C., Switaia J., Crowley D., Conley M.:
RA "The delta (argP-lacZ)205(0169) deletion greatly enhances resistance
RT to hydrogen peroxide in stationary-phase *Escherichia coli*.";
RL J. Bacteriol. 176:1297-1302(1994).
RN [5]
RP IDENTIFICATION.
RA Rudd K.B.:
RL Unpublished observations (APR-1994).
CC -!- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 257.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000134; AAC73370.1; -;
DR EMBL; D83536; BAA77933.1; -;
DR EMBL; U70214; AAB08688.1; -;
DR EMBL; L20943; -; NOT ANNOTATED_CDS.
DR EcoGene; EG12338; yaGA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 261 271 MISSING (IN REF. 1).
SQ SEQUENCE 384 AA; 43768 MW; 96619AF7C82FED5E CRC64;
Query Match 91.2%; Score 31; DB 1; Length 384;
Best Local Similarity 80.0%; Pred. No. 42;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CVAHC 5
DB 175 CLAHC 179
RESULT 2
YDZ5 SCHPO STANDARD; PRT; 844 AA.
AC 013712;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C14C4.05c in chromosome I.
GN SPAC14C4.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC
RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsis K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Ruster S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.:
RT "The genome sequence of *Schizosaccharomyces pombe*.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: SOME TO YEAST YML034W.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z98596; CAB11199.1; -;
DR EMBL; Z98596; CAB11199.1; -;
KW Hypothetical protein; Transmembrane; ATP-binding.
FT TRANSMEM 459 479 POTENTIAL.
FT TRANSMEM 705 725 POTENTIAL.
FT NP_BIND 605 612 ATP (POTENTIAL).
SQ SEQUENCE 844 AA; 94907 MW; D31C398350014766 CRC64;
Query Match 91.2%; Score 31; DB 1; Length 844;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CVAHC 5
DB 745 CVSHC 749
RESULT 3
C110 HUMAN STANDARD; PRT; 1069 AA.
ID C110 HUMAN
AC Q9NZB2; Q9NZB1; Q14688;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein C9orf10.
GN C9orf10 OR KIAA0183.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RP Brahmbhatt S.B., Hulme D.J., Dawkins J.L., Nicholson G.A.:
RA "Generating full-length coding sequence for 2 alternate transcripts of
RT a novel gene C9orf10 and screening HSN-1 patients for mutations.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE OF 8-1069 FROM N.A. (ISOFORM A).

OX NCBI_TaxID=6221;

RN [1]
 RP SEQUENCE.
 RX MEDLINE=81264350; PubMed=7263698;
 RA Blumenthal K.M., Keim P.S., Heinrichson R.L., Kem W.R.;
 RT "Structure and action of heteronemertine polypeptide toxins. Amino
 RT acid sequence of Cerebratulus lacteus toxin B-II and revised
 RT structure of toxin B-IV.";
 RL J. Biol. Chem. 256:9063-9067(1981).
 CC -1- SIMILARITY: TO NEUROTOXIN B-IV.
 DR PIR; A01789; NTHB2.
 DR HSP; P01525; IIVB.
 KW Toxin; Hydroxylation.
 FT MOD RES 10 10 HYDROXYLATION.
 FT DISULFID 12 48 BY SIMILARITY.
 FT DISULFID 16 52 BY SIMILARITY.
 FT DISULFID 23 41 BY SIMILARITY.
 FT DISULFID 26 37 BY SIMILARITY.
 SQ SEQUENCE 55 AA; 6028 MW; 0918C1DFB2FA904 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 55;
 Best Local Similarity 80.0%; Pred. No. 11;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVAHC 5
 |
 |
 |
 |
 Db 37 CAABC 41

RESULT 7

NXB4 CERLA
 ID NXB4 CERLA STANDARD; PRT; 55 AA.
 AC P01525;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Neurotoxin B-IV.
 OS Cerebratulus lacteus (Milky ribbon worm).
 OC Eukaryota; Metazoa; Nemertea; Anopla; Heteronemertea; Cerebratulus.
 OX NCBI_TaxID=6221;
 RN [1]
 RP SEQUENCE, AND REVISIONS.
 RX MEDLINE=81264350; PubMed=7263698;
 RA Blumenthal K.M., Keim P.S., Heinrichson R.L., Kem W.R.;
 RT "Structure and action of heteronemertine polypeptide toxins. Amino
 RT acid sequence of Cerebratulus lacteus toxin B-II and revised
 RT structure of toxin B-IV.";
 RL J. Biol. Chem. 256:9063-9067(1981).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=77006160; PubMed=972152;
 RA Blumenthal K.M., Kem W.R.;
 RT "Structure and action of heteronemertine polypeptide toxins. Primary
 RT structure of Cerebratulus lacteus toxin B-IV.";
 RL J. Biol. Chem. 251:6025-6029(1976).
 RN [3]
 RP MUTAGENESIS.
 RX MEDLINE=91302300; PubMed=2071577;
 RA Howell M.L., Blumenthal K.M.;
 RT "Mutagenesis of Cerebratulus lacteus neurotoxin B-IV identifies NH2-
 RT terminal sequences important for biological activity.";
 RL J. Biol. Chem. 266:12884-12888(1991).
 RN [4]
 RP STRUCTURE BY NMR.
 RX MEDLINE=93076792; PubMed=1332861;
 RA Hansen P.E., Kem W.R., Bieber A.L., Norton R.S.;
 RT "1H-NMR study of neurotoxin B-IV from the marine worm Cerebratulus
 RT lacteus. Solution properties, sequence-specific resonance
 RT assignments, secondary structure and global fold.";
 RL Eur. J. Biochem. 210:231-240(1992).
 RN [5]
 RP STRUCTURE BY NMR.
 RX MEDLINE=97324210; PubMed=9180379;

RA Barnham K.J., Dyke T.R., Kem W.R., Norton R.S.;
 RT "Structure of neurotoxin B-IV from the marine worm Cerebratulus
 RT lacteus: a helical hairpin cross-linked by disulphide bonding.";
 RL J. Mol. Biol. 268:886-902(1997).
 CC -1- FUNCTION: ONLY ACTS ON SOME CRUSTACEAN. INCREASE THE EXCITABILITY
 CC OF NERVES PROBABLY BY AFFECTING THE INACTIVATION OF THE VOLTAGE-
 CC GATED SODIUM CHANNEL.
 CC -1- SIMILARITY: TO NEUROTOXIN B-II.
 DR PIR; A01788; NTHB4.
 DR PDB; 1VIB; 15-MAY-97.
 KW Toxin; Hydroxylation; 3D-structure.
 FT MOD RES 10 10 HYDROXYLATION.
 FT DISULFID 12 48
 FT DISULFID 16 52
 FT DISULFID 23 41
 FT DISULFID 26 37
 SQ SEQUENCE 55 AA; 6107 MW; BB76B72E48DB050D CRC64;

Query Match 88.2%; Score 30; DB 1; Length 55;
 Best Local Similarity 80.0%; Pred. No. 11;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVAHC 5
 |
 |
 |
 |
 Db 37 CAABC 41

RESULT 8

DEF1 STOCA
 ID DEF1 STOCA STANDARD; PRT; 79 AA.
 AC O16136;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Defensin 1 precursor.
 GN SMD1.
 OS Stomoxys calcitrans (Stable fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Muscoidea; Muscidae; Stomoxys.
 OX NCBI_TaxID=35570;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Midgut;
 RX MEDLINE=97470996; PubMed=9326639;
 RA Lehane M.J., Wu D., Lehane S.M.;
 RT "Midgut-specific immune molecules are produced by the blood-sucking
 RT insect Stomoxys calcitrans.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11502-11507(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Midgut;
 RA Munks R.J.L., Lehane S.M., Lehane M.J.;
 RT "Expression and regulation of the genes, smd1 and smd2, which encode
 RT midgut-specific defensin molecules in Stomoxys calcitrans.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF013146; AAB66892.1; --
 CC EMBL; AF182163; AAD56535.1; --
 CC HSP; P10891; IICA.
 DR InterPro; IPR001542; Defensin_anopod.
 DR InterPro; IPR003614; Knot1.
 DR Pfam; PF01097; Arthro defensin; 1.
 DR PRINTS; PR00271; DEFENSIN.

DR SMART; SM00505; Knot1; 1.
 DR PROSITE; PS00425; ARTHROPOD DEFENSINS; 1.
 KW Insect immunity; Antibiotic; Signal.
 FT SIGNAL 1 23
 FT PROPEP 24 ?

FT CHAIN ?
 FT DISULFID 42 69
 FT DISULFID 55 75
 FT DISULFID 59 77
 SQ SEQUENCE 79 AA; 8160 MW; 5F275B39EB641778 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 79;
 Best Local Similarity 80.0%; Pred. No. 16;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 Db 55 CAAC 59

RESULT 9

DEFI_TENMO STANDARD; PRT; 84 AA.
 AC Q27023;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Tenebrion 1 precursor.
 OS Tenebrio molitor (Yellow mealworm).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Cucujiformia; Tenebrionidae; Tenebrio.
 OX NCBI_TaxID=7067;
 RN [1]

SEQUENCE FROM N.A., SEQUENCE OF 49-59, AND DISULFIDE BONDS.

TISSUE=Larval hemolymph;
 RX MEDLINE=95096025; PubMed=7798186;
 RA Moon H.J., Lee S.Y., Kurata S., Natori S., Lee B.L.;
 RT "Purification and molecular cloning of cDNA for an inducible
 RT antibacterial protein from larvae of the coleopteran, Tenebrio
 RT molitor.";
 RL J. Biochem. 116:53-58(1994).
 CC -1- FUNCTION: BACTERICIDAL PROTEIN PRODUCED IN RESPONSE TO INJURY. IT
 CC IS CYTOTOXIC TO GRAM-POSITIVE BACTERIA.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; D17670; BAA04552.1; --
 DR HSPF; P10891; IICA.
 DR InterPro; IPR001542; Defensin_annot.
 DR InterPro; IPR003614; Knot1.
 DR Pfam; PF01097; Arthro defensin; 1.
 DR PRINTS; PR00271; DEFENSIN.
 DR SMART; SM00505; Knot1; 1.

DR PROSITE; PS00425; ARTHROPOD DEFENSINS; 1.
 KW Insect immunity; Antibiotic; Signal.
 FT SIGNAL 1 19
 FT PROPEP 20 41
 FT CHAIN 42 84
 FT DISULFID 44 75
 FT DISULFID 61 81
 FT DISULFID 65 83

SEQUENCE 84 AA; 9176 MW; 0367C5070468BE60 CRC64;
 88.2%; Score 30; DB 1; Length 84;

ch 88.2%; Score 30; DB 1; Length 84;

Best Local Similarity 80.0%; Pred. No. 17;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 Db 61 CAAC 65

RESULT 10

DEFI_PROTE STANDARD; PRT; 94 AA.
 AC P10891;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Phormicin precursor (Insect defensins A and B).
 OS Protophormia terrae novae (Black blowfly) (Nestling-sucking blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Oestroidea; Calliphoridae; Protophormia.
 OX NCBI_TaxID=34676;
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=90316108; PubMed=2369900;
 RA Dimarcq J.-L., Zachary D., Hoffmann J.A., Hoffmann D.,
 RA Reichhart J.-M.;
 RT "Insect immunity: expression of the two major inducible antibacterial
 RT peptides, defensin and dipterocin, in Phormia terrae novae.";
 RL EMBO J. 9:2507-2515(1990).
 RN [2]

SEQUENCE OF 55-94.

TISSUE=Hemolymph;
 RX MEDLINE=89098894; PubMed=2911573;
 RA Lambert J., Keppi E., Dimarcq J.-L., Wicker C., Reichhart J.-M.,
 RA Dunbar B., Lepage P., van Dorsselaer A., Hoffmann J.A., Fothergill J.,
 RA Hoffmann D.;
 RT "Insect immunity: isolation from immune blood of the dipteran Phormia
 RT terrae novae of two insect antibacterial peptides with sequence
 RT homology to rabbit lung macrophage bactericidal peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:262-266(1989).
 RN [3]

DISULFIDE BONDS.

RX MEDLINE=91192047; PubMed=2013293;
 RA Lepage P., Bitsch P., Roeklin D., Keppi E., Dimarcq J.-L.,
 RA Reichhart J.-M., Hoffmann J.A., Roitsch C., van Dorsselaer A.;
 RT "Determination of disulfide bridges in natural and recombinant insect
 RT defensin A.";
 RL Eur. J. Biochem. 196:735-742(1991).
 RN [4]

STRUCTURE BY NMR.

RX MEDLINE=93104264; PubMed=1467342;
 RA Bonmatin J.M., Genest M., Petit M.-C., Labbe H., Vovelle F., Ptak M.,
 RA Cornet B., Gallet X., Callet A., Labbe H., Vovelle F., Ptak M.;
 RT "Progress in multidimensional NMR investigations of peptide and
 RT protein 3-D structures in solution. From structure to functional
 RT aspects.";
 RL Biochimie 74:825-836(1992).
 RN [5]

STRUCTURE BY NMR.

RX MEDLINE=95393015; PubMed=7663941;
 RA Cornet B., Bonmatin J.M., Hetru C., Hoffmann J.A., Ptak M.,
 RA Vovelle F.;

RT "Refined three-dimensional solution structure of insect defensin A.";
 RL Structure 3:435-448(1995).

CC -1- FUNCTION: RESPONSIBLE FOR THE ANTI GRAM-POSITIVE ACTIVITY OF
 CC IMMUNE HEMOLYMPH OF P.TERRAE NOVAE.

CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X55546; CAA39152.1; --

DR PIR; A32219; A32219.

DR PIR; B32219; B32219.

DR PIR; S12558; S12558.

DR PDB; 1ICA; 10-JUL-95.

DR InterPro; IPR001542; Defensin_anod.

DR InterPro; IPR003614; Knot1.

DR Pfam; PF01097; Arthro_defensin; 1.

DR PRINTS; PR00271; DEFENSIN.

DR SMART; SM00505; Knot1; 1.

DR PROSITE; PS00425; ARTHROPOD DEFENSINS; 1.

KW Insect immunity; Antibiotic; Signal; 3D-structure.

FT SIGNAL 1 23

FT PROPEP 24 54

FT CHAIN 55 94

FT DISULFID 57 84

FT DISULFID 70 90

FT DISULFID 74 92

FT VARIANT 86 86

FT SEQUENCE 94 AA; 10110 MW; CE24ED83D7CB86D9 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 94;

Best Local Similarity 80.0%; Pred. No. 18;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CVAHC 5

Db 70 CAABC 74

RESULT 11

ID SAPE_SARPE STANDARD; PRT; 94 AA.

AC P18313;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Sapacin precursor.

OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Oestroidea; Sarcophagidae; Sarcophaga.

OX NCBI_TaxId=7386;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89034216; PubMed=3182837;

RA Matsuyama K., Natori S.;

RT "Molecular cloning of cDNA for sapacin and unique expression of the sapacin gene during the development of Sarcophaga peregrina.";

RL J. Biol. Chem. 263:17117-17121(1988).

RN [2]

RP SEQUENCE OF 55-94.

RX MEDLINE=89034215; PubMed=3182836;

RA Matsuyama K., Natori S.;

RT "Purification of three antibacterial proteins from the culture medium of NIH-Sape-4, an embryonic cell line of Sarcophaga peregrina.";

RL J. Biol. Chem. 263:17112-17116(1988).

RN [3]

RP DISULFIDE BONDS.

RX MEDLINE=90292974; PubMed=2358424;

RA Kuzuhara T., Nakajima Y., Matsuyama K., Natori S.;

RT "Determination of the disulfide array in sapacin, an antibacterial peptide of Sarcophaga peregrina (flesh fly).";

RL J. Biochem. 107:514-518(1990).

RN [4]

RP STRUCTURE BY NMR.

RX MEDLINE=90382590; PubMed=2401368;

RA Hanzawa H., Shimada I., Kuzuhara T., Komano H., Kohda D., Inagaki F., Natori S., Arata Y.;

RT "1H nuclear magnetic resonance study of the solution conformation of

an antibacterial protein, sapacin.";

FFBS Lett. 269:413-420(1990).

-1- FUNCTION: SAPECIN, WHICH ARE POTENT BACTERICIDAL PROTEINS, ARE PRODUCED IN RESPONSE TO INJURY. SAPECIN IS CYTOTOXIC TO GRAM-POSITIVE BACTERIA, AND IN A LESSER EXTENT AGAINST GRAM-NEGATIVE BACTERIA.

-1- TISSUE SPECIFICITY: HEMOCYTES AND FAT BODY.

-1- INDUCTION: BY INJURY TO THE LARVAL CELL WALL.

-1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; J04053; AAA29984.1; --

DR PIR; A31791; A31791.

DR PIR; A31792; A31792.

DR HSP; P10891; 1ICA.

DR InterPro; IPR001542; Defensin_anod.

DR InterPro; IPR003614; Knot1.

DR Pfam; PF01097; Arthro_defensin; 1.

DR PRINTS; PR00271; DEFENSIN.

DR SMART; SM00505; Knot1; 1.

DR PROSITE; PS00425; ARTHROPOD DEFENSINS; 1.

KW Insect immunity; Antibiotic; Cleavage on pair of basic residues;

KW Signal.

FT SIGNAL 1 23

FT PROPEP 24 54

FT PEPTIDE 55 94

FT DISULFID 57 84

FT DISULFID 70 90

FT DISULFID 74 92

FT SEQUENCE 94 AA; 9914 MW; 36170AB3073EC947 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 94;

Best Local Similarity 80.0%; Pred. No. 18;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CVAHC 5

Db 70 CAABC 74

RESULT 12

ID DEF2_STOCA STANDARD; PRT; 97 AA.

AC O16137;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Defensin 2 precursor.

GN SMD2.

OS Stomoxys calcitrans (Stable fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Muscoidea; Muscidae; Stomoxys.

OX NCBI_TaxId=35570;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 58-67.

RC TISSUE=Midgut;

RX MEDLINE=97470996; PubMed=9326639;

RA Lehane M.J., Wu D., Lehane S.M.;

RT "Midgut-specific immune molecules are produced by the blood-sucking insect Stomoxys calcitrans.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:11502-11507(1997).

-1- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE MOSTLY AGAINST GRAM-NEGATIVE BACTERIA. ITS ACTIVITY IS ENHANCED BY LIPOPOLYSACCHARIDE.

-1- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN THE ANTERIOR

FT	CHAIN	60	99	DEFENSIN C.
FT	DISULFID	62	89	BY SIMILARITY.
FT	DISULFID	75	95	BY SIMILARITY.
FT	DISULFID	79	97	BY SIMILARITY.
FT	VARIANT	26	26	G -> E (IN ISOFORM C2).
FT	VARIANT	34	34	P -> S (IN ISOFORM C2).
FT	CONFLICT	83	83	R -> G (IN REF. 2 AND 3).
FT	CONFLICT	91	91	A -> S (IN REF. 2 AND 3).
SSQ	SEQUENCE	99 AA;	10696 MW;	7C3566A68F4548ED CRC64;

Query Match	88.2%;	Score 30;	DB 1;	Length 99;
Best Local Similarity	80.0%;	Pred. NO. 19;		
Matches	4;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;

OY	1	CVANC 5
DDb	75	CAAHC 79

RESULT 14	?			
DEFI_ANOGA				
IID	DEFI_ANOGA	STANDARD;	PRT;	102 AA.
AC	Q17027;			
DDT	01-NOV-1997	(Rel. 35, Created)		
DDT	01-NOV-1997	(Rel. 35, Last sequence update)		
DDT	01-NOV-1997	(Rel. 35, Last annotation update)		
DE	Defensin precursor.			
OS	Anopheles gambiae (African malaria mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;			
OC	Culicoides; Anopheles.			
OX	NCHI_TaxID=7165;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RRP	STRAIN=G3;			
RC	Richmann A. M.			
RA				

RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RESPONSIBLE FOR THE ANTI GRAM-POSITIVE ACTIVITY OF
CC IMMUNE HEMOLYMPH (BY SIMILARITY).
CC
CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X93562; CAA63775.1; ALT_INIT.
DR HSP; P10891; IICA.
DR InterPro; IPR001542; Defensin_anpod.
DR InterPro; IPR003614; Knot1.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PRINTS; PR00271; DEFENSIN.
DR SMART; SM00505; Knot1. 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Insect immunity; Antibiotic; Signal.
FT SIGNAL 1 25
FT PROPEP 26 62
FT CHAIN 63 102
FT DISULFID 65 92
FT DISULFID 78 98
FT DISULFID 82 100
SQ SEQUENCE 102 AA; 10627 MW; 628834560AEDCD0C CRC64;

Query Match 88.2%; Score 30; DB 1; Length 102;
Best Local Similarity 80.0%; Pred. No. 20;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVAHC 5
DB 78 CAABC 82

RESULT 15
Y304 METJA
ID Y304 METJA STANDARD; PRT; 159 AA.
AC Q5752;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ0304.
GN MJ0304.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=9637999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervase A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R.; Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073 (1996).
CC -1- SIMILARITY: BELONGS TO THE GAMMA-CLASS CARBONIC ANHYDRASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U67485; AAB98291.1;
DR HSP; P40881; IQRE.
DR TIGR; M0304;
DR InterPro; IPR001451; Hexapep_transf.
DR Pfam; PF00132; hexapep; 5.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 159 AA; 17174 MW; 8EBB072E411BCC94 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 159;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVAHC 5
DB 51 CVVHC 55

RESULT 16
XDRB_ECO57
ID XDRB_ECO57 STANDARD; PRT; 292 AA.
AC Q8X6C5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Xanthine dehydrogenase, FAD binding subunit (EC 1.1.1.204).
GN XDRB OR Z4206 OR ECS3740.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258786;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtaubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogawara N., Yasunaga T.,
RA Kunihara S., Shiba T., Hattori M., Shinagawa H.,
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22 (2001).
RN [1]
RP FUNCTION: Presumed to be a dehydrogenase, but possibly an oxidase.
CC Participates in limited purine salvage (requires aspartate) but
CC does not support aerobic growth on purines as the sole carbon
CC source [purine catabolism] (by similarity).
CC -1- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
CC -1- COFACTOR: FAD (by similarity).
CC -1- PATHWAY: Purine catabolism; first committed step.
CC -1- SUBUNIT: Heterotrimer of xdhA, xdhB and xdhC (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

```
CC EMBL; AE005516; AAG57996.1; -
DR EMBL; AP002563; BAB37163.1; -
KW Oxidoreductase; NAD; Flavoprotein; FAD; Purine metabolism;
KW Purine salvage; Complete proteome.
SQ SEQUENCE 292 AA; 31561 MW; FEC44F9980BF9BC1 CRC64;

Query Match      88.2%; Score 30; DB 1; Length 292;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CVAHC 5 /
Db 199 CAHC 203

RESULT 17
XDBH_ECOLI
ID XDBH_ECOLI STANDARD; PRT; 292 AA.
AC Q46800;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Xanthine dehydrogenase, FAD binding subunit (EC 1.1.1.204).
GN XDBH OR B2867.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III; Bloch C.A.; Perna N.T.; Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP DISCUSSION OF FUNCTION.
RC STRAIN=K12 / W3110;
RX MEDLINE=20444178; PubMed=10986234;
RA Xi H., Schneider B.L., Reitzer L.;
RT "Purine catabolism in Escherichia coli and function of xanthine
dehydrogenase in purine salvage.";
RL J. Bacteriol. 182:5332-5341(2000).
CC -!- FUNCTION: Presumed to be a dehydrogenase, but possibly an oxidase.
CC Participates in limited purine salvage (requires aspartate) but
CC does not support aerobic growth on purines as the sole carbon
CC source (purine catabolism).
CC -!- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
CC -!- COFACTOR: FAD (By similarity).
CC -!- PATHWAY: Purine catabolism; first committed step.
CC -!- SUBUNIT: Heterotrimer of xdhA, xdhB and xdhC (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28375; AAA83048.1; -
DR EMBL; AE000370; AAC75905.1; -
DR EcoGene; EG13050; xdhB
DR InterPro; IPR005107; CO deh flav C.
DR InterPro; IPR002346; dehydrog_molyb.
KW PF00941; FAD binding 5; 1.
KW PF03450; CO deh flav C; 1.
KW Phosphatase; NAD; Flavoprotein; FAD; Purine metabolism;
KW Complete proteome.
```

```
SQ SEQUENCE 292 AA; 31556 MW; 74A07D137DA857B8 CRC64;

Query Match      88.2%; Score 30; DB 1; Length 292;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CVAHC 5
Db 199 CAHC 203

RESULT 18
GRHR_MOUSE
ID GRHR_MOUSE STANDARD; PRT; 327 AA.
AC Q0176; Q61611;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gonadotropin-releasing hormone receptor (GNRH-R).
GN GNRHR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92375111; PubMed=1324422;
RX Tautumi M., Zhou W., Millar R.P., Mellon P.L., Roberts J.L.,
RA Flanagan C.A., Dong K., Gillo B., Sealton S.C.;
RA "Cloning and functional expression of a mouse gonadotropin-releasing
hormone receptor.";
RT Mol. Endocrinol. 6:1163-1169(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=93016057; PubMed=1328228;
RX Reinhardt J., Mertz L.M., Catt K.J.;
RT "Molecular cloning and expression of cDNA encoding the murine
gonadotropin-releasing hormone receptor.";
RL J. Biol. Chem. 267:21281-21284(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Clay C.M., Nelson S.E., Campion C.E., Digregorio G.B.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH
CC G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: PITUITARY.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L01119; AAB59636.1; -
DR EMBL; M93108; -; NOT ANNOTATED_CDS.
DR EMBL; L33789; AAA37716.1; -
DR EMBL; L33778; AAA37716.1; JOINED.
DR EMBL; L33788; AAA37716.1; JOINED.
DR PIR; A45101; A45101.
DR MGD; MGI:95790; Gnrhr.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G PROTEIN RECEPT FL 1; 1.
DR PROSITE; PS00262; G PROTEIN RECEPT FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 39 58 1 (POTENTIAL).
FT
```

```
FT DOMAIN 59 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 97 2 (POTENTIAL).
FT DOMAIN 98 115 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 116 137 3 (POTENTIAL).
FT DOMAIN 138 164 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 165 184 4 (POTENTIAL).
FT DOMAIN 185 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 231 5 (POTENTIAL).
FT DOMAIN 232 280 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 281 299 6 (POTENTIAL).
FT DOMAIN 300 305 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 306 325 7 (POTENTIAL).
FT DOMAIN 326 327 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 114 195 BY SIMILARITY.
FT CONFLICT 76 76 M -> I (IN REF. 3).
SQ SEQUENCE 327 AA; 37684 MW; FDDC7B985306FC9F CRC64;

Query Match 88.2%; Score 30; DB 1; Length 327;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 195 CVTHC 199

RESULT 19
GRHR RAT
ID GRHR RAT STANDARD; PRT; 327 AA.
AC P30969;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Gonadotropin-releasing hormone receptor (GNRH-R).
GN GNRHR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Pituitary;
RX MEDLINE=93246029; PubMed=1338727;
RA Eidine K.A., Sella R.E., Couper G., Anderson L., Taylor P.L.;
RT "Molecular cloning and characterization of the rat pituitary
gonadotropin-releasing hormone (GNRH) receptor.";
RL Mol. Cell. Endocrinol. 90:R5-R9(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Buffalo; TISSUE=Pituitary;
RA Kakar S.S., Grantham K., Musgrove L.C., Devor D.C., Sellers J.C.,
RA Neill J.D.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=93221478; PubMed=7916600;
RA Perrin M.H., Bilezikjian L.M., Hoeger C., Donaldson C.J., Rivier J.,
RA Haas Y., Vale W.W.;
RT "Molecular and functional characterization of GNRH receptors cloned
from rat pituitary and a mouse pituitary tumor cell line.";
RL Biochem. Biophys. Res. Commun. 191:1139-1144(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Pituitary;
RX MEDLINE=93129239; PubMed=1339279;
RA Kaiser U.B., Zhao D., Cardona G.R., Chin W.W.;
RT "Isolation and characterization of cDNAs encoding the rat pituitary
gonadotropin-releasing hormone receptor.";
RL Biochem. Biophys. Res. Commun. 189:1645-1652(1992).
```

```
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Gonad;
RX MEDLINE=94241995; PubMed=8185587;
RA Mounni M., Kottler M.L., Counis R.;
RT "Nucleotide sequence analysis of mRNAs predicts that rat pituitary
and gonadal gonadotropin-releasing hormone receptor proteins have
identical primary structure.";
RL Biochem. Biophys. Res. Commun. 200:1359-1366(1994).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=94122663; PubMed=7764374;
RA Kudo A., Park M.K., Kawashima S.;
RT "Isolation of rat GnRH receptor cDNA having different 5'-noncoding
sequence.";
RL Zool. Sci. 10:863-867(1993).
RN [7]
RP SEQUENCE FROM N.A.
RA Reinhart J., Xiao S., Arora K.K., Catt K.J.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH
CC G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69880; CAA48776.1; -
CC EMBL; S59525; AAB28420.1; -
CC EMBL; U00935; AAC27349.1; -
CC EMBL; L25053; AAA41265.1; ALT_INIT.
CC EMBL; L07646; AAA41274.1; -
CC EMBL; X76635; CAA54083.1; -
CC EMBL; S68578; AAC60671.1; -
CC EMBL; U92471; AAB58038.1; -
CC EMBL; U92469; AAB58038.1; JOINED.
CC EMBL; U92470; AAB58038.1; JOINED.
CC PIR; S29332; S29332.
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC Pfam: PF00001; 7cm.1; 1.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 39 58 1 (POTENTIAL).
CC DOMAIN 59 77 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 78 97 2 (POTENTIAL).
CC DOMAIN 98 115 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 116 137 3 (POTENTIAL).
CC DOMAIN 138 164 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 165 184 4 (POTENTIAL).
CC DOMAIN 185 211 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 212 231 5 (POTENTIAL).
CC DOMAIN 232 280 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 281 299 6 (POTENTIAL).
CC DOMAIN 300 305 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 306 325 7 (POTENTIAL).
CC DOMAIN 326 327 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC DISULFID 114 195 BY SIMILARITY.
CC CONFLICT 184 184 A -> V (IN REF. 1).
CC CONFLICT 272 272 G -> A (IN REF. 3, 5, 6 AND 7).
CC CONFLICT 311 311 A -> G (IN REF. 1).
CC SEQUENCE 327 AA; 37748 MW; 86A2CA1A2C9F1BE4 CRC64;
```


Query Match 88.2%; Score 30; DB 1; Length 327;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 195 CVTHC 199

RESULT 20

GRHR_BOVIN STANDARD; PRT; 328 AA.
ID GRHR_BOVIN STANDARD; PRT; 328 AA.
AC P32236;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gonadotropin-releasing hormone receptor (GNRH-R).
GN GNRH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94139286; PubMed=8306635;
RA Kakar S.S., Rane C.H., Neill J.D.;
RT "Molecular cloning, sequencing, and characterizing the bovine
RT receptor for gonadotropin releasing hormone (GNRH).";
RL Domest. Anim. Endocrinol. 10:335-342(1993).
RN [2]
RP SEQUENCE OF 1-74 FROM N.A.
RA Nelson S.E., Quirk C.E., Clay C.M.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH
CC G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U00934; AAC48857.1; --
CC EMBL; AF034950; AAB91470.1; --
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN RECEPTOR F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein.
CC DOMAIN 1 38
CC FT EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 39 58
CC FT DOMAIN 59 77
CC FT TRANSMEM 78 97
CC FT TRANSMEM 98 115
CC FT TRANSMEM 116 137
CC FT TRANSMEM 138 164
CC FT TRANSMEM 165 184
CC FT TRANSMEM 185 212
CC FT TRANSMEM 213 232
CC FT TRANSMEM 233 281
CC FT TRANSMEM 282 300
CC FT TRANSMEM 301 306
CC FT TRANSMEM 307 326
CC FT TRANSMEM 327 328
CC FT TRANSMEM 18 18
CC FT TRANSMEM 102 102
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DISULFID 114 196 BY SIMILARITY.
SQ SEQUENCE 328 AA; 37740 MW; 845312BAB53B4F72 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 328;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 196 CVTHC 200

RESULT 21

GRHR_HORSE STANDARD; PRT; 328 AA.
ID GRHR_HORSE STANDARD; PRT; 328 AA.
AC O18921;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadotropin-releasing hormone receptor (GNRH-R).
GN GNRH.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX Porter M.B., Green M.L., Simmen R.C.M., Sharp D.C.;
RA "Cloning and sequencing of the equine pituitary gonadotropin-releasing
RA hormone receptor and expression of its mRNA following administration
RA of pulseatile or continuous GnRH.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH
CC G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF018072; AAC27099.2; --
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN RECEPTOR F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein.
CC DOMAIN 1 38
CC FT EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 39 59
CC FT DOMAIN 60 84
CC FT TRANSMEM 85 105
CC FT DOMAIN 106 115
CC FT TRANSMEM 116 136
CC FT DOMAIN 137 157
CC FT TRANSMEM 158 178
CC FT DOMAIN 179 208
CC FT TRANSMEM 209 229
CC FT DOMAIN 230 271
CC FT TRANSMEM 272 292
CC FT DOMAIN 293 306
CC FT TRANSMEM 307 327
CC FT DOMAIN 328 328
CC FT CARBOHYD 18 18
CC FT DISULFID 114 196
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC ADD7E25A41EACD0A CRC64;
SQ SEQUENCE 328 AA; 37648 MW; 88.2%; Score 30; DB 1; Length 328;

Query Match

SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=96235149; PubMed=8666259;
Campion C.E., Turzillo A.M., Clay C.M.;
"The gene encoding the ovine gonadotropin-releasing hormone (GnRH)
receptor: cloning and initial characterization.";
Gene 170:277-280(1996).
-1- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH
G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
MESSENGER SYSTEM.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; L22215; AAC37336.1; -
EMBL; X72088; CAA50978.1; -
EMBL; L42937; AAB38515.1; -
EMBL; L43842; AAB41939.1; -
EMBL; L43841; AAB41939.1; JOINED.
PIR; JN0882; JN0882.
InterPro; IPR000276; GPCR_Rhodopsn.
Pfam; PF00001; 7tm_1; 1.
PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
G-protein coupled receptor; Transmembrane; Glycoprotein.
DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
TRANSMEM 39 58 1 (POTENTIAL).
DOMAIN 59 77 CYTOPLASMIC (POTENTIAL).
TRANSMEM 78 97 2 (POTENTIAL).
DOMAIN 98 115 EXTRACELLULAR (POTENTIAL).
TRANSMEM 116 137 3 (POTENTIAL).
DOMAIN 138 164 CYTOPLASMIC (POTENTIAL).
TRANSMEM 165 184 4 (POTENTIAL).
DOMAIN 185 212 EXTRACELLULAR (POTENTIAL).
TRANSMEM 213 232 5 (POTENTIAL).
DOMAIN 233 281 CYTOPLASMIC (POTENTIAL).
TRANSMEM 282 300 6 (POTENTIAL).
DOMAIN 301 306 EXTRACELLULAR (POTENTIAL).
TRANSMEM 307 326 7 (POTENTIAL).
DOMAIN 327 328 CYTOPLASMIC (POTENTIAL).
TRANSMEM 329 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
DOMAIN 331 332 BY SIMILARITY.
TRANSMEM 333 334 S -> R (IN REF. 2).
DOMAIN 335 336 S -> A (IN REF. 2).
TRANSMEM 337 338 A -> G (IN REF. 2).
DOMAIN 339 340
TRANSMEM 341 342
DOMAIN 343 344
TRANSMEM 345 346
DOMAIN 347 348
TRANSMEM 349 350
DOMAIN 351 352
TRANSMEM 353 354
DOMAIN 355 356
TRANSMEM 357 358
DOMAIN 359 360
TRANSMEM 361 362
DOMAIN 363 364
TRANSMEM 365 366
DOMAIN 367 368
TRANSMEM 369 370
DOMAIN 371 372
TRANSMEM 373 374
DOMAIN 375 376
TRANSMEM 377 378
DOMAIN 379 380
TRANSMEM 381 382
DOMAIN 383 384
TRANSMEM 385 386
DOMAIN 387 388
TRANSMEM 389 390
DOMAIN 391 392
TRANSMEM 393 394
DOMAIN 395 396
TRANSMEM 397 398
DOMAIN 399 400
TRANSMEM 401 402
DOMAIN 403 404
TRANSMEM 405 406
DOMAIN 407 408
TRANSMEM 409 410
DOMAIN 411 412
TRANSMEM 413 414
DOMAIN 415 416
TRANSMEM 417 418
DOMAIN 419 420
TRANSMEM 421 422
DOMAIN 423 424
TRANSMEM 425 426
DOMAIN 427 428
TRANSMEM 429 430
DOMAIN 431 432
TRANSMEM 433 434
DOMAIN 435 436
TRANSMEM 437 438
DOMAIN 439 440
TRANSMEM 441 442
DOMAIN 443 444
TRANSMEM 445 446
DOMAIN 447 448
TRANSMEM 449 450
DOMAIN 451 452
TRANSMEM 453 454
DOMAIN 455 456
TRANSMEM 457 458
DOMAIN 459 460
TRANSMEM 461 462
DOMAIN 463 464
TRANSMEM 465 466
DOMAIN 467 468
TRANSMEM 469 470
DOMAIN 471 472
TRANSMEM 473 474
DOMAIN 475 476
TRANSMEM 477 478
DOMAIN 479 480
TRANSMEM 481 482
DOMAIN 483 484
TRANSMEM 485 486
DOMAIN 487 488
TRANSMEM 489 490
DOMAIN 491 492
TRANSMEM 493 494
DOMAIN 495 496
TRANSMEM 497 498
DOMAIN 499 500
TRANSMEM 501 502
DOMAIN 503 504
TRANSMEM 505 506
DOMAIN 507 508
TRANSMEM 509 510
DOMAIN 511 512
TRANSMEM 513 514
DOMAIN 515 516
TRANSMEM 517 518
DOMAIN 519 520
TRANSMEM 521 522
DOMAIN 523 524
TRANSMEM 525 526
DOMAIN 527 528
TRANSMEM 529 530
DOMAIN 531 532
TRANSMEM 533 534
DOMAIN 535 536
TRANSMEM 537 538
DOMAIN 539 540
TRANSMEM 541 542
DOMAIN 543 544
TRANSMEM 545 546
DOMAIN 547 548
TRANSMEM 549 550
DOMAIN 551 552
TRANSMEM 553 554
DOMAIN 555 556
TRANSMEM 557 558
DOMAIN 559 560
TRANSMEM 561 562
DOMAIN 563 564
TRANSMEM 565 566
DOMAIN 567 568
TRANSMEM 569 570
DOMAIN 571 572
TRANSMEM 573 574
DOMAIN 575 576
TRANSMEM 577 578
DOMAIN 579 580
TRANSMEM 581 582
DOMAIN 583 584
TRANSMEM 585 586
DOMAIN 587 588
TRANSMEM 589 590
DOMAIN 591 592
TRANSMEM 593 594
DOMAIN 595 596
TRANSMEM 597 598
DOMAIN 599 600
TRANSMEM 601 602
DOMAIN 603 604
TRANSMEM 605 606
DOMAIN 607 608
TRANSMEM 609 610
DOMAIN 611 612
TRANSMEM 613 614
DOMAIN 615 616
TRANSMEM 617 618
DOMAIN 619 620
TRANSMEM 621 622
DOMAIN 623 624
TRANSMEM 625 626
DOMAIN 627 628
TRANSMEM 629 630
DOMAIN 631 632
TRANSMEM 633 634
DOMAIN 635 636
TRANSMEM 637 638
DOMAIN 639 640
TRANSMEM 641 642
DOMAIN 643 644
TRANSMEM 645 646
DOMAIN 647 648
TRANSMEM 649 650
DOMAIN 651 652
TRANSMEM 653 654
DOMAIN 655 656
TRANSMEM 657 658
DOMAIN 659 660
TRANSMEM 661 662
DOMAIN 663 664
TRANSMEM 665 666
DOMAIN 667 668
TRANSMEM 669 670
DOMAIN 671 672
TRANSMEM 673 674
DOMAIN 675 676
TRANSMEM 677 678
DOMAIN 679 680
TRANSMEM 681 682
DOMAIN 683 684
TRANSMEM 685 686
DOMAIN 687 688
TRANSMEM 689 690
DOMAIN 691 692
TRANSMEM 693 694
DOMAIN 695 696
TRANSMEM 697 698
DOMAIN 699 700
TRANSMEM 701 702
DOMAIN 703 704
TRANSMEM 705 706
DOMAIN 707 708
TRANSMEM 709 710
DOMAIN 711 712
TRANSMEM 713 714
DOMAIN 715 716
TRANSMEM 717 718
DOMAIN 719 720
TRANSMEM 721 722
DOMAIN 723 724
TRANSMEM 725 726
DOMAIN 727 728
TRANSMEM 729 730
DOMAIN 731 732
TRANSMEM 733 734
DOMAIN 735 736
TRANSMEM 737 738
DOMAIN 739 740
TRANSMEM 741 742
DOMAIN 743 744
TRANSMEM 745 746
DOMAIN 747 748
TRANSMEM 749 750
DOMAIN 751 752
TRANSMEM 753 754
DOMAIN 755 756
TRANSMEM 757 758
DOMAIN 759 760
TRANSMEM 761 762
DOMAIN 763 764
TRANSMEM 765 766
DOMAIN 767 768
TRANSMEM 769 770
DOMAIN 771 772
TRANSMEM 773 774
DOMAIN 775 776
TRANSMEM 777 778
DOMAIN 779 780
TRANSMEM 781 782
DOMAIN 783 784
TRANSMEM 785 786
DOMAIN 787 788
TRANSMEM 789 790
DOMAIN 791 792
TRANSMEM 793 794
DOMAIN 795 796
TRANSMEM 797 798
DOMAIN 799 800
TRANSMEM 801 802
DOMAIN 803 804
TRANSMEM 805 806
DOMAIN 807 808
TRANSMEM 809 810
DOMAIN 811 812
TRANSMEM 813 814
DOMAIN 815 816
TRANSMEM 817 818
DOMAIN 819 820
TRANSMEM 821 822
DOMAIN 823 824
TRANSMEM 825 826
DOMAIN 827 828
TRANSMEM 829 830
DOMAIN 831 832
TRANSMEM 833 834
DOMAIN 835 836
TRANSMEM 837 838
DOMAIN 839 840
TRANSMEM 841 842
DOMAIN 843 844
TRANSMEM 845 846
DOMAIN 847 848
TRANSMEM 849 850
DOMAIN 851 852
TRANSMEM 853 854
DOMAIN 855 856
TRANSMEM 857 858
DOMAIN 859 860
TRANSMEM 861 862
DOMAIN 863 864
TRANSMEM 865 866
DOMAIN 867 868
TRANSMEM 869 870
DOMAIN 871 872
TRANSMEM 873 874
DOMAIN 875 876
TRANSMEM 877 878
DOMAIN 879 880
TRANSMEM 881 882
DOMAIN 883 884
TRANSMEM 885 886
DOMAIN 887 888
TRANSMEM 889 890
DOMAIN 891 892
TRANSMEM 893 894
DOMAIN 895 896
TRANSMEM 897 898
DOMAIN 899 900
TRANSMEM 901 902
DOMAIN 903 904
TRANSMEM 905 906
DOMAIN 907 908
TRANSMEM 909 910
DOMAIN 911 912
TRANSMEM 913 914
DOMAIN 915 916
TRANSMEM 917 918
DOMAIN 919 920
TRANSMEM 921 922
DOMAIN 923 924
TRANSMEM 925 926
DOMAIN 927 928
TRANSMEM 929 930
DOMAIN 931 932
TRANSMEM 933 934
DOMAIN 935 936
TRANSMEM 937 938
DOMAIN 939 940
TRANSMEM 941 942
DOMAIN 943 944
TRANSMEM 945 946
DOMAIN 947 948
TRANSMEM 949 950
DOMAIN 951 952
TRANSMEM 953 954
DOMAIN 955 956
TRANSMEM 957 958
DOMAIN 959 960
TRANSMEM 961 962
DOMAIN 963 964
TRANSMEM 965 966
DOMAIN 967 968
TRANSMEM 969 970
DOMAIN 971 972
TRANSMEM 973 974
DOMAIN 975 976
TRANSMEM 977 978
DOMAIN 979 980
TRANSMEM 981 982
DOMAIN 983 984
TRANSMEM 985 986
DOMAIN 987 988
TRANSMEM 989 990
DOMAIN 991 992
TRANSMEM 993 994
DOMAIN 995 996
TRANSMEM 997 998
DOMAIN 999 1000
TRANSMEM 1001 1002
DOMAIN 1003 1004
TRANSMEM 1005 1006
DOMAIN 1007 1008
TRANSMEM 1009 1010
DOMAIN 1011 1012
TRANSMEM 1013 1014
DOMAIN 1015 1016
TRANSMEM 1017 1018
DOMAIN 1019 1020
TRANSMEM 1021 1022
DOMAIN 1023 1024
TRANSMEM 1025 1026
DOMAIN 1027 1028
TRANSMEM 1029 1030
DOMAIN 1031 1032
TRANSMEM 1033 1034
DOMAIN 1035 1036
TRANSMEM 1037 1038
DOMAIN 1039 1040
TRANSMEM 1041 1042
DOMAIN 1043 1044
TRANSMEM 1045 1046
DOMAIN 1047 1048
TRANSMEM 1049 1050
DOMAIN 1051 1052
TRANSMEM 1053 1054
DOMAIN 1055 1056
TRANSMEM 1057 1058
DOMAIN 1059 1060
TRANSMEM 1061 1062
DOMAIN 1063 1064
TRANSMEM 1065 1066
DOMAIN 1067 1068
TRANSMEM 1069 1070
DOMAIN 1071 1072
TRANSMEM 1073 1074
DOMAIN 1075 1076
TRANSMEM 1077 1078
DOMAIN 1079 1080
TRANSMEM 1081 1082
DOMAIN 1083 1084
TRANSMEM 1085 1086
DOMAIN 1087 1088
TRANSMEM 1089 1090
DOMAIN 1091 1092
TRANSMEM 1093 1094
DOMAIN 1095 1096
TRANSMEM 1097 1098
DOMAIN 1099 1100
TRANSMEM 1101 1102
DOMAIN 1103 1104
TRANSMEM 1105 1106
DOMAIN 1107 1108
TRANSMEM 1109 1110
DOMAIN 1111 1112
TRANSMEM 1113 1114
DOMAIN 1115 1116
TRANSMEM 1117 1118
DOMAIN 1119 1120
TRANSMEM 1121 1122
DOMAIN 1123 1124
TRANSMEM 1125 1126
DOMAIN 1127 1128
TRANSMEM 1129 1130
DOMAIN 1131 1132
TRANSMEM 1133 1134
DOMAIN 1135 1136
TRANSMEM 1137 1138
DOMAIN 1139 1140
TRANSMEM 1141 1142
DOMAIN 1143 1144
TRANSMEM 1145 1146
DOMAIN 1147 1148
TRANSMEM 1149 1150
DOMAIN 1151 1152
TRANSMEM 1153 1154
DOMAIN 1155 1156
TRANSMEM 1157 1158
DOMAIN 1159 1160
TRANSMEM 1161 1162
DOMAIN 1163 1164
TRANSMEM 1165 1166
DOMAIN 1167 1168
TRANSMEM 1169 1170
DOMAIN 1171 1172
TRANSMEM 1173 1174
DOMAIN 1175 1176
TRANSMEM 1177 1178
DOMAIN 1179 1180
TRANSMEM 1181 1182
DOMAIN 1183 1184
TRANSMEM 1185 1186
DOMAIN 1187 1188
TRANSMEM 1189 1190
DOMAIN 1191 1192
TRANSMEM 1193 1194
DOMAIN 1195 1196
TRANSMEM 1197 1198
DOMAIN 1199 1200
TRANSMEM 1201 1202
DOMAIN 1203 1204
TRANSMEM 1205 1206
DOMAIN 1207 1208
TRANSMEM 1209 1210
DOMAIN 1211 1212
TRANSMEM 1213 1214
DOMAIN 1215 1216
TRANSMEM 1217 1218
DOMAIN 1219 1220
TRANSMEM 1221 1222
DOMAIN 1223 1224
TRANSMEM 1225 1226
DOMAIN 1227 1228
TRANSMEM 1229 1230
DOMAIN 1231 1232
TRANSMEM 1233 1234
DOMAIN 1235 1236
TRANSMEM 1237 1238
DOMAIN 1239 1240
TRANSMEM 1241 1242
DOMAIN 1243 1244
TRANSMEM 1245 1246
DOMAIN 1247 1248
TRANSMEM 1249 1250
DOMAIN 1251 1252
TRANSMEM 1253 1254
DOMAIN 1255 1256
TRANSMEM 1257 1258
DOMAIN 1259 1260
TRANSMEM 1261 1262
DOMAIN 1263 1264
TRANSMEM 1265 1266
DOMAIN 1267 1268
TRANSMEM 1269 1270
DOMAIN 1271 1272
TRANSMEM 1273 1274
DOMAIN 1275 1276
TRANSMEM 1277 1278
DOMAIN 1279 1280
TRANSMEM 1281 1282
DOMAIN 1283 1284
TRANSMEM 1285 1286
DOMAIN 1287 1288
TRANSMEM 1289 1290
DOMAIN 1291 1292
TRANSMEM 1293 1294
DOMAIN 1295 1296
TRANSMEM 1297 1298
DOMAIN 1299 1300
TRANSMEM 1301 1302
DOMAIN 1303 1304
TRANSMEM 1305 1306
DOMAIN 1307 1308
TRANSMEM 1309 1310
DOMAIN 1311 1312
TRANSMEM 1313 1314
DOMAIN 1315 1316
TRANSMEM 1317 1318
DOMAIN 1319 1320
TRANSMEM 1321 1322
DOMAIN 1323 1324
TRANSMEM 1325 1326
DOMAIN 1327 1328
TRANSMEM 1329 1330
DOMAIN 1331 1332
TRANSMEM 1333 1334
DOMAIN 1335 1336
TRANSMEM 1337 1338
DOMAIN 1339 1340
TRANSMEM 1341 1342
DOMAIN 1343 1344
TRANSMEM 1345 1346
DOMAIN 1347 1348
TRANSMEM 1349 1350
DOMAIN 1351 1352
TRANSMEM 1353 1354
DOMAIN 1355 1356
TRANSMEM 1357 1358
DOMAIN 1359 1360
TRANSMEM 1361 1362
DOMAIN 1363 1364
TRANSMEM 1365 1366
DOMAIN 1367 1368
TRANSMEM 1369 1370
DOMAIN 1371 1372
TRANSMEM 1373 1374
DOMAIN 1375 1376
TRANSMEM 1377 1378
DOMAIN 1379 1380
TRANSMEM 1381 1382
DOMAIN 1383 1384
TRANSMEM 1385 1386
DOMAIN 1387 1388
TRANSMEM 1389 1390
DOMAIN 1391 1392
TRANSMEM 1393 1394
DOMAIN 1395 1396
TRANSMEM 1397 1398
DOMAIN 1399 1400
TRANSMEM 1401 1402
DOMAIN 1403 1404
TRANSMEM 1405 1406
DOMAIN 1407 1408
TRANSMEM 1409 1410
DOMAIN 1411 1412
TRANSMEM 1413 1414
DOMAIN 1415 1416
TRANSMEM 1417 1418
DOMAIN 1419 1420
TRANSMEM 1421 1422
DOMAIN 1423 1424
TRANSMEM 1425 1426
DOMAIN 1427 1428
TRANSMEM 1429 1430
DOMAIN 1431 1432
TRANSMEM 1433 1434
DOMAIN 1435 1436
TRANSMEM 1437 1438
DOMAIN 1439 1440
TRANSMEM 1441 1442
DOMAIN 1443 1444
TRANSMEM 1445 1446
DOMAIN 1447 1448
TRANSMEM 1449 1450
DOMAIN 1451 1452
TRANSMEM 1453 1454
DOMAIN 1455 1456
TRANSMEM 1457 1458
DOMAIN 1459 1460
TRANSMEM 1461 1462
DOMAIN 1463 1464
TRANSMEM 1465 1466
DOMAIN 1467 1468
TRANSMEM 1469 1470
DOMAIN 1471 1472
TRANSMEM 1473 1474
DOMAIN 1475 1476
TRANSMEM 1477 1478
DOMAIN 1479 1480
TRANSMEM 1481 1482
DOMAIN 1483 1484
TRANSMEM 1485 1486
DOMAIN 1487 1488
TRANSMEM 1489 1490
DOMAIN 1491 1492
TRANSMEM 1493 1494
DOMAIN 1495 1496
TRANSMEM 1497 1498
DOMAIN 1499 1500
TRANSMEM 1501 1502
DOMAIN 1503 1504
TRANSMEM 1505 1506
DOMAIN 1507 1508
TRANSMEM 1509 1510
DOMAIN 1511 1512
TRANSMEM 1513 1514
DOMAIN 1515 1516
TRANSMEM 1517 1518
DOMAIN 1519 1520
TRANSMEM 1521 1522
DOMAIN 1523 1524
TRANSMEM 1525 1526
DOMAIN 1527 1528
TRANSMEM 1529 1530
DOMAIN 1531 1532
TRANSMEM 1533 1534
DOMAIN 1535 1536
TRANSMEM 1537 1538
DOMAIN 1539 1540
TRANSMEM 1541 1542
DOMAIN 1543 1544
TRANSMEM 1545 1546
DOMAIN 1547 1548
TRANSMEM 1549 1550
DOMAIN 1551 1552
TRANSMEM 1553 1554
DOMAIN 1555 1556
TRANSMEM 1557 1558
DOMAIN 1559 1560
TRANSMEM 1561 1562
DOMAIN 1563 1564
TRANSMEM 1565 1566
DOMAIN 1567 1568
TRANSMEM 1569 1570
DOMAIN 1571 1572
TRANSMEM 1573 1574
DOMAIN 1575 1576
TRANSMEM 1577 1578
DOMAIN 1579 1580
TRANSMEM 1581 1582
DOMAIN 1583 1584
TRANSMEM 1585 1586
DOMAIN 1587 1588
TRANSMEM 1589 1590
DOMAIN 1591 1592
TRANSMEM 1593 1594
DOMAIN 1595 1596
TRANSMEM 1597 1598
DOMAIN 1599 1600
TRANSMEM 1601 1602
DOMAIN 1603 1604
TRANSMEM 1605 1606
DOMAIN 1607 1608
TRANSMEM 1609 1610
DOMAIN 1611 1612
TRANSMEM 1613 1614
DOMAIN 1615 1616
TRANSMEM 1617 1618
DOMAIN 1619 1620
TRANSMEM 1621 1622
DOMAIN 1623 1624
TRANSMEM 1625 1626
DOMAIN 1627 1628
TRANSMEM 1629 1630
DOMAIN 1631 1632
TRANSMEM 1633 1634
DOMAIN 1635 1636
TRANSMEM 1637 1638
DOMAIN 1639 1640
TRANSMEM 1641 1642
DOMAIN 1643 1644
TRANSMEM 1645 1646
DOMAIN 1647 1648
TRANSMEM 1649 1650
DOMAIN 1651 1652
TRANSMEM 1653 1654
DOMAIN 1655 1656
TRANSMEM 1657 1658
DOMAIN 1659 1660
TRANSMEM 1661 1662
DOMAIN 1663 1664
TRANSMEM 1665 1666
DOMAIN 1667 1668
TRANSMEM 1669 1670
DOMAIN 1671 1672
TRANSMEM 1673 1674
DOMAIN 1675 1676
TRANSMEM 1677 1678
DOMAIN 1679 1680
TRANSMEM 1681 1682
DOMAIN 1683 1684
TRANSMEM 1685 1686
DOMAIN 1687 1688
TRANSMEM 1689 1690
DOMAIN 1691 1692
TRANSMEM 1693 1694
DOMAIN 1695 1696
TRANSMEM 1697 1698
DOMAIN 1699 1700
TRANSMEM 1701 1702
DOMAIN 1703 1704
TRANSMEM 1705 1706
DOMAIN 1707 1708
TRANSMEM 1709 1710
DOMAIN 1711 1712
TRANSMEM 1713 1714
DOMAIN 1715 1716
TRANSMEM 1717 1718
DOMAIN 1719 1720
TRANSMEM 1721 1722
DOMAIN 1723 1724
TRANSMEM 1725 1726
DOMAIN 1727 1728
TRANSMEM 1729 1730
DOMAIN 1731 1732
TRANSMEM 1733 1734
DOMAIN 1735 1736
TRANSMEM 1737 1738
DOMAIN 1739 1740
TRANSMEM 1741 1742
DOMAIN 1743 1744
TRANSMEM 1745 1746
DOMAIN 1747 1748
TRANSMEM 1749 1750
DOMAIN 1751 1752
TRANSMEM 1753 1754
DOMAIN 1755 1756
TRANSMEM 1757 1758
DOMAIN 1759 1760
TRANSMEM 1761 1762
DOMAIN 1763 1764
TRANSMEM 1765 1766
DOMAIN 1767 1768
TRANSMEM 1769 1770
DOMAIN 1771 1772
TRANSMEM 1773 1774
DOMAIN 1775 1776
TRANSMEM 1777 1778
DOMAIN 1779 1780
TRANSMEM 1781 1782
DOMAIN 1783 1784
TRANSMEM 1785 1786
DOMAIN 1787 1788
TRANSMEM 1789 1790
DOMAIN 1791 1792
TRANSMEM 1793 1794
DOMAIN 1795 1796
TRANSMEM 1797 1798
DOMAIN 1799 1800
TRANSMEM 1801 1802
DOMAIN 1803 1804
TRANSMEM 1805 1806
DOMAIN 1807 1808
TRANSMEM 1809 1810
DOMAIN 1811 1812
TRANSMEM 1813 1814
DOMAIN 1815 1816
TRANSMEM 1817 1818
DOMAIN 1819 1820
TRANSMEM 1821 1822
DOMAIN 1823 1824
TRANSMEM 1825 1826
DOMAIN 1827 1828
TRANSMEM 1829 1830
DOMAIN 1831 1832
TRANSMEM 1833 1834
DOMAIN 1835 1836
TRANSMEM 1837 1838
DOMAIN 1839 1840
TRANSMEM 1841 1842
DOMAIN 1843 1844
TRANSMEM 1845 1846
DOMAIN 1847 1848
TRANSMEM 1849 1850
DOMAIN 1851 1852
TRANSMEM 1853 1854
DOMAIN 1855 1856
TRANSMEM 1857 1858
DOMAIN 1859 1860
TRANSMEM 1861 1862
DOMAIN 1863 1864
TRANSMEM 1865 1866
DOMAIN 1867 1868
TRANSMEM 1869 1870
DOMAIN 1871 1872
TRANSMEM 1873 1874
DOMAIN 1875 1876
TRANSMEM 1877 1878
DOMAIN 1879 1880
TRANSMEM 1881 1882
DOMAIN 1883 1884
TRANSMEM 1885 1886
DOMAIN 1887 1888
TRANSMEM 1889 1890
DOMAIN 1891 1892
TRANSMEM 1893 1894
DOMAIN 1895 1896
TRANSMEM 1897 1898
DOMAIN 1899 1900
TRANSMEM 1901 1902
DOMAIN 1903 1904
TRANSMEM 1905 1906
DOMAIN 1907 1908
TRANSMEM 1909 1910
DOMAIN 1911 1912
TRANSMEM 1913 1914
DOMAIN 1915 1916
TRANSMEM 1917 1918
DOMAIN 1919 1920
TRANSMEM 1921 1922
DOMAIN 1923 1924
TRANSMEM 1925 1926
DOMAIN 1927 1928
TRANSMEM 1929 1930
DOMAIN 1931 1932
TRANSMEM 1933 1934
DOMAIN 1935 1936
TRANSMEM 1937 1938
DOMAIN 1939 1940
TRANSMEM 1941 1942
DOMAIN 1943 1944
TRANSMEM 1945 1946
DOMAIN 1947 1948
TRANSMEM 1949 1950
DOMAIN 1951 1952
TRANSMEM 1953 1954
DOMAIN 1955 1956
TRANSMEM 1957 1958
DOMAIN 1959 1960
TRANSMEM 1961 1962
DOMAIN 1963 1964
TRANSMEM 1965 1966
DOMAIN 1967 1968
TRANSMEM 1969 1970
DOMAIN 1971 1972
TRANSMEM 1973 1974
DOMAIN 1975 1976
TRANSMEM 1977 1978
DOMAIN 1979 1980
TRANSMEM 1981 1982
DOMAIN 1983 1984
TRANSMEM 1985 1986
DOMAIN 1987 1988
TRANSMEM 1989 1990
DOMAIN 1991 1992
TRANSMEM 1993 1994
DOMAIN 1995 1996
TRANSMEM 1997 1998
DOMAIN 1999 2000
TRANSMEM 2001 2002
DOMAIN 2003 2004
TRANSMEM 2005 2006
DOMAIN 2007 2008
TRANSMEM 2009 2010
DOMAIN 2011 2012
TRANSMEM 2013 2014
DOMAIN 2015 2016
TRANSMEM 2017 2018
DOMAIN 2019 2020
TRANSMEM 2021 2022
DOMAIN 2023 2024
TRANSMEM 2025 2026
DOMAIN 2027 2028
TRANSMEM 2029 2030
DOMAIN 2

OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20433225; PubMed=10977898;
RA Strych U., Huang H.-C., Krause K.L., Benedik M.J.;
RT "Characterization of the alanine racemases from Pseudomonas aeruginosa
RL PAOI";
RN Curr. Microbiol. 41:290-294 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Coltray L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sajer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen";
RL Nature 406:959-964 (2000).
CC -!- FUNCTION: Isomerizes L-alanine to D-alanine which is then oxidized
CC to pyruvate by dadA (By similarity).
CC -!- CATALYTIC ACTIVITY: L-alanine = D-alanine.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Alanine catabolism; first step.
CC -!- INDUCTION: By alanine.
CC -!- SIMILARITY: BELONGS TO THE ALANINE RACEMASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF165881; AAD47081.1; -.
DR EMBL; AE004943; AAG08687.1; -.
DR HSP; P10724; 1BD0.
DR InterPro; IPR000821; Ala_racemase.
DR Pfam; PF00842; Ala_racemase; 1.
DR PRINTS; PR00992; ALARACEMASE.
DR TIGRFAMs; TIGR00492; alt; 1.
DR PROSITE; PS00395; ALANINE RACEMASE; 1.
KW Isomerase; Pyridoxal phosphate; Complete proteome.
FT ACT_SITE 33 33 CATALYTIC BASE SPECIFIC TO D-ALANINE (BY
FT ACT_SITE 253 253 CATALYTIC BASE SPECIFIC TO L-ALANINE (BY
FT ACT_SITE 33 33 SIMILARITY).
FT BINDING 33 33 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT CONFLICT 138 138 S -> R (IN REF. 1).
FT CONFLICT 178 178 A -> S (IN REF. 1).
FT CONFLICT 262 262 S -> R (IN REF. 1).
SQ SEQUENCE 357 AA; 38914 MW; 199F4021D02FF48 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 357;
Best Local Similarity: 80.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 97 CVAHC 101

Search completed: June '20, 2003, 20:53:02
Job time : 8.02273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:42:49 ; Search time 26.3636 Seconds
(without alignment)
39.078 Million cell updates/sec

Title: US-10-105-008-18
Perfect score: 34
Sequence: 1 CVAHC 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SPTREMBL 21:*

- 1: sp archaea:*
- 2: sp bacteria:*
- 3: sp fungi:*
- 4: sp human:*
- 5: sp invertebrate:*
- 6: sp mammal:*
- 7: sp mhc:*
- 8: sp organelle:*
- 9: sp phage:*
- 10: sp plant:*
- 11: sp rodent:*
- 12: sp virus:*
- 13: sp vertebrate:*
- 14: sp unclassified:*
- 15: sp rvirus:*
- 16: sp bacteriaph:*
- 17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	67	16 Q9JZM3	Q9jzm3 neisseria m
2	34	100.0	87	5 Q8WTD4	Q8wtd4 glossina mo
3	34	100.0	305	16 Q67090	Q67090 aquifex aeo
4	34	100.0	438	16 Q9PEM7	Q9pem7 xylella fas
5	34	100.0	540	5 Q9Y1A5	Q9y1a5 lymnaea sta
6	34	100.0	547	2 Q07454	Q07454 rhodospseudo
7	34	100.0	967	5 Q9Y1A6	Q9y1a6 lymnaea sta
8	33	97.1	144	4 Q96SH9	Q96sh9 homo sapien
9	33	97.1	166	4 Q96LV5	Q96lv5 mus musculu
10	33	97.1	166	11 Q9CQJ8	Q9cuj8 mus musculu
11	31	91.2	88	5 P91335	P91335 caenorhabdi
12	31	91.2	133	10 Q9FUG3	Q9fug3 typha latif
13	31	91.2	138	12 Q8QRA1	Q8qra1 avian infec
14	31	91.2	141	11 Q9JJ96	Q9jj96 mus musculu
15	31	91.2	141	12 Q8QRA0	Q8qra0 avian infec
16	31	91.2	173	5 O44981	O44981 caenorhabdi

RESULT 1

ID	Q9JZM3	PRELIMINARY;	PRT;	67 AA.
AC	Q9JZM3			
DT	01-OCT-2000	(TEMBLrel. 15, Created)		
DT	01-OCT-2000	(TEMBLrel. 15, Last sequence update)		
DT	01-DEC-2001	(TEMBLrel. 19, Last annotation update)		
DE	Hypothetical protein NMB0988.			
GN	NMB0988.			

ALIGNMENTS

17	31	91.2	182	16	O66549	O66549 aquifex aeo
18	31	91.2	199	3	Q9USD6	Q9usd6 schizosacch
19	31	91.2	298	4	Q9NWX6	Q9nwx6 homo sapien
20	31	91.2	298	4	Q9HOS2	Q9hos2 homo sapien
21	31	91.2	340	16	Q9X7W1	Q9x7w1 streptomyce
22	31	91.2	400	16	Q55546	Q55546 synchocyst.
23	31	91.2	414	5	Q20629	Q20629 caenorhabdi
24	31	91.2	443	10	Q9RT40	Q9rt40 arabidopsis
25	31	91.2	451	10	Q9CA63	Q9ca63 arabidopsis
26	31	91.2	485	10	Q9ZPM2	Q9zpm2 arabidopsis
27	31	91.2	485	10	Q9M6R0	Q9m6r0 arabidopsis
28	31	91.2	485	10	Q941P0	Q941p0 brassica ju
29	31	91.2	485	10	Q8W122	Q8w122 arabidopsis
30	31	91.2	485	10	Q9S7Z3	Q9s7z3 arabidopsis
31	31	91.2	498	10	Q8SAG8	Q8sag8 glycine max
32	31	91.2	500	10	Q9SWM5	Q9swm5 triticum ae
33	31	91.2	502	10	Q9XHP8	Q9xhp8 arabidopsis
34	31	91.2	676	5	P91957	P91957 musca domes
35	31	91.2	697	11	Q91VX9	Q91vx9 mus musculu
36	31	91.2	715	2	Q9X5N7	Q9x5n7 escherichia
37	31	91.2	851	2	O51812	O51812 agrobacteri
38	31	91.2	1481	5	Q8SRQ1	Q8srq1 encephalito
39	30	88.2	39	5	P82380	P82380 stomoxys ca
40	30	88.2	57	5	Q9GYU6	Q9gyu6 aedes albop
41	30	88.2	63	4	Q9NNX2	Q9nnx2 homo sapien
42	30	88.2	72	5	Q9BLM0	Q9blm0 androctonus
43	30	88.2	75	12	Q64945	Q64945 avian infec
44	30	88.2	75	12	Q64951	Q64951 avian infec
45	30	88.2	75	12	Q64952	Q64952 avian infec
46	30	88.2	75	12	Q64953	Q64953 avian infec
47	30	88.2	76	12	Q64943	Q64943 avian infec
48	30	88.2	76	12	Q64944	Q64944 avian infec
49	30	88.2	76	12	Q64946	Q64946 avian infec
50	30	88.2	76	12	Q64947	Q64947 avian infec
51	30	88.2	76	12	Q64948	Q64948 avian infec
52	30	88.2	76	12	Q64949	Q64949 avian infec
53	30	88.2	76	12	Q64950	Q64950 avian infec
54	30	88.2	79	5	Q96049	Q96049 oryctes rhi
55	30	88.2	79	5	P82378	P82378 stomoxys ca
56	30	88.2	81	4	O75812	O75812 homo sapien
57	30	88.2	96	5	O77217	O77217 aedes albop
58	30	88.2	97	5	P82379	P82379 stomoxys ca
59	30	88.2	98	5	Q963E9	Q963e9 aedes aegypt
60	30	88.2	98	5	P91793	P91793 aedes aegypt
61	30	88.2	102	5	O61721	O61721 anopheles g
62	30	88.2	102	11	Q9D3S1	Q9d3s1 mus musculu
63	30	88.2	123	12	Q9WML2	Q9wml2 avian infec
64	30	88.2	128	11	Q925Q5	Q925q5 mus musculu
65	30	88.2	129	4	Q96T91	Q96t91 homo sapien
66	30	88.2	130	11	Q925Q4	Q925q4 rattus norv
67	30	88.2	130	12	Q8Q888	Q8q888 avian infec
68	30	88.2	133	12	Q91SA9	Q91sa9 avian infec
69	30	88.2	133	12	Q8QRA2	Q8qra2 avian infec
70	30	88.2	133	12	Q8Q930	Q8qg93 avian infec
71	30	88.2	133	12	Q8QQ90	Q8qq90 avian infec
72	30	88.2	133	12	Q8Q887	Q8q887 avian infec
73	30	88.2	134	12	Q8QRA4	Q8qra4 avian infec
74	30	88.2	136	12	Q91SB0	Q91sb0 avian infec
75	30	88.2	137	12	Q8Q891	Q8q891 avian infec

OS *Neisseria meningitidis* (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491,
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
RT MC58";
RL Science 287:1809-1815 (2000).
DR EMBL; AF002449; AAF41391.1; -.
DR TIGR; NMB0988; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 67 AA; 7296 MW; D093968E392A719F CRC64;

Query Match 100.0%; Score 34; DB 16; Length 67;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
Db 9 CVAHC 13

RESULT 2

OSWTD4
ID Q8WTD4 PRELIMINARY; PRT; 87 AA.
AC Q8WTD4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Antimicrobial peptide defensin DefA.
OS *Glossina morsitans morsitans*.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Hippoboscidae; Glossinidae; Glossina.
OX NCBI_TaxID=37546;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21532909; PubMed=11592981;
RA Hao Z., Kasumba I., Lehane M.J., Gibson W.C., Kwon J., Aksoy S.;
RT "Tsetse immune responses and trypanosome transmission: implications
RT for the development of tsetse-based strategies to reduce
RT trypanosomiasis";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12648-12653 (2001).
DR EMBL; AF368907; AAL34112.1; -.
SQ SEQUENCE 87 AA; 9592 MW; C160DF6FAC8E5413 CRC64;

Query Match 100.0%; Score 34; DB 5; Length 87;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
Db 64 CVAHC 68

RESULT 3

OSWTD4
ID Q67090 PRELIMINARY; PRT; 305 AA.
AC Q67090;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE PHOSPHOTIDYLGLYCEROPHOSPHATE synthase.

GN PGSA2 OR AQ_958.
OS *Aquifex aeolicus*.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.B., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium *Aquifex
RT aeolicus*";
RL Nature 392:353-358 (1998).
DR EMBL; AE000715; AAC07048.1; -.
DR InterPro; IPR000462; CDP-OH_P_transf.
DR InterPro; IPR004570; PGSA.
DR Pfam; PF01066; CDP-OH_P_transf; 1.
DR TIGRPFAMs; TIGR00560; PGSA; 1.
DR PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; 1.
KW Complete proteome.
SQ SEQUENCE 305 AA; 35265 MW; 60054265BF996723 CRC64;

Query Match 100.0%; Score 34; DB 16; Length 305;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
Db 296 CVAHC 300

RESULT 4

OSPEM7
ID Q9PEM7 PRELIMINARY; PRT; 438 AA.
AC Q9PEM7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Acetylglutamate kinase.
GN XF1001.
OS *Xylella fastidiosa*.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Agencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S.; Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite J.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri B.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Besquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawaaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva P.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."
RL Nature 406:151-159(2000).
DR EMBL; AE003938; AAF83811.1; -.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR004662; Agglutininase.
DR Pfam; PF00696; aakinas; 1.
DR TIGRFAMs; TIGR00761; argB; 1.
KW Complete proteome.
SQ SEQUENCE 438 AA; 48939 MW; D64F22FEM4617781 CRC64;

Query Match 100.0%; Score 34; DB 16; Length 438;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 425 CVAHC 429

RESULT 5

ID Q9Y1A5 PRELIMINARY; PRT; 540 AA.
AC Q9Y1A5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 13, Last annotation update)
DE Furinl (Fragment).
GN FUR1.
OS *Lymnaea stagnalis* (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBRAL GANGLION;
RX MEDLINE=99456930; PubMed=10526318;
RA Spijker S., Smit A.B., Sharp-Baker H.E., Van Elk R.,
RA Van Kesteren E.R., Van Minnen J., Kurosky A., Geraerts W.P.;
RT "Family of prohormone convertases in *Lymnaea*: characterization of two
RT alternatively spliced furin-like transcripts and cell-specific
RT regulation of their expression."
RL J. Neurobiol. 41:399-413(1999).
DR EMBL; AF140362; AAD44729.1; -.
DR MEROPS; S08.071; -.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR002884; P_domain.
DR InterPro; IPR002994; VitK_dep_GLA.
DR Pfam; PF01483; P; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR ProDom; PD000717; P_domain; 1.
DR SMART; SM00261; FU; 2.
DR PROSITE; PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
SQ SEQUENCE 540 AA; 58790 MW; F6F355699EC399F8 CRC64;

Query Match 100.0%; Score 34; DB 5; Length 540;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 362 CVAHC 366

PRELIMINARY; PRT; 547 AA.
TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cyclohex-1-ene-1-carboxylate CoA ligase.
GN ALIA.
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009;
RX MEDLINE=97322399; PubMed=9177244;
RA Eglund P.G., Pelletier D.A., Dispensa M., Gibson J., Harwood C.S.;
RT "A cluster of bacterial genes for anaerobic benzene ring
RT biodegradation."
RL Proc. Natl. Acad. Sci. U.S.A. 94:6484-6489(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009;
RX MEDLINE=98233726; PubMed=9573182;
RA Pelletier D.A., Harwood C.S.;
RT "2-Ketocyclohexanecarboxyl coenzyme A hydrolase, the ring cleavage
RT enzyme required for anaerobic benzoate degradation by Rhodopseudomonas
RT palustris."
RL J. Bacteriol. 180:2330-2336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009;
RA Eglund P.G., Pelletier D.A., Dispensa M., Gibson J., Harwood C.S.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U75363; AAC23919.1; -.
DR HSP; P08659; ILCI.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Ligase.
SQ SEQUENCE 547 AA; 59507 MW; 0A210A3D57AC1713 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 33 CVAHC 37

RESULT 7

ID Q9Y1A6 PRELIMINARY; PRT; 967 AA.
AC Q9Y1A6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Furinl-X.
GN FUR1.
OS *Lymnaea stagnalis* (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBRAL GANGLION;
RX MEDLINE=99456930; PubMed=10526318;
RA Spijker S., Smit A.B., Sharp-Baker H.E., Van Elk R.,
RA Van Kesteren E.R., Van Minnen J., Kurosky A., Geraerts W.P.;
RT "Family of prohormone convertases in *Lymnaea*: characterization of two
RT alternatively spliced furin-like transcripts and cell-specific
RT regulation of their expression."
RL J. Neurobiol. 41:399-413(1999).
DR EMBL; AF140361; AAD44728.1; -.
DR HSP; Q99405; IMPT.
DR MEROPS; S08.071; -.

```

DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR002884; P_domain.
DR InterPro; IPR000294; Vtk_dep_GLA.
DR Pfam; PF01483; P; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD000717; P_domain; 1.
DR SMART; SM00261; FU; 2.
DR PROSITE; PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
SQ SEQUENCE 967 AA; 107600 MW; 85EA0988D1B964A4 CRC64;

Query Match 100.0%; Score 34; DB 5; Length 967;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 789 CVAHC 793

RESULT 8
Q96SH9 PRELIMINARY; PRT; 144 AA.
AC Q96SH9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DJ113701.1 (Similar to putative mitochondrial inner membrane protease
DE subunit 2) (Fragment).
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Frankland J.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132295; CAC39221.1; -.
DR MEROFS; S26.013; -.
DR InterPro; IPR000508; SigPTase.
DR Pfam; PF00461; Peptidase_S26; 1.
DR PROSITE; PS00761; SPASE_I_3; UNKNOWN_1.
FT NON TER 144
SQ SEQUENCE 144 AA; 15987 MW; 9FBD0F560D9E3A2 CRC64;

Query Match 97.1%; Score 33; DB 4; Length 144;
Best Local Similarity 80.0%; Pred. No. 21;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 21 CVAHC 25

RESULT 9
Q96LU5 PRELIMINARY; PRT; 166 AA.
AC Q96LU5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDNA FLJ25059 f18, clone CBL04610.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=CEREBELLUM;
RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK057788; BAB71573.1; -.
DR InterPro; IPR000508; SigPTase.
DR Pfam; PF00461; Peptidase_S26; 1.
DR PROSITE; PS00761; SPASE_I_3; UNKNOWN_1.
SQ SEQUENCE 166 AA; 18504 MW; 53E4EB65757BA71A CRC64;

Query Match 97.1%; Score 33; DB 4; Length 166;
Best Local Similarity 80.0%; Pred. No. 24;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVAHC 5
DB 21 CVAHC 25

RESULT 10
Q9CQU8 PRELIMINARY; PRT; 166 AA.
AC Q9CQU8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 1500034J20RIK protein (2610528017RIK protein) (RIKEN cDNA 1500034J20
DE gene).
GN 1500034J20RIK OR 2610528017RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CS7BL/6J; TISSUE=TESTIS, CEREBELLUM, TONGUE, AND EMBRYO;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L. M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boriss A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK015978; BAB30062.1; -.
DR EMBL; AK005356; BAB23971.1; -.
DR EMBL; AK009570; BAB26367.1; -.
DR EMBL; AK010075; BAB26684.1; -.
DR EMBL; AK011278; BAB27512.1; -.
DR EMBL; AK011382; BAB27582.1; -.
DR EMBL; AK011420; BAB27608.1; -.
DR EMBL; AK011466; BAB27637.1; -.

```

DR EMBL; AK011618; BAB27737.1; -;
 DR EMBL; AK011872; BAB27891.1; -;
 DR EMBL; AK012181; BAB28082.1; -;
 DR EMBL; BC008259; AAH08259.1; -;
 DR HSP; P00803; I812.
 DR MEROPS; S26.013; -;
 DR MGD; MGI:1913791; 1500034J20R1k.
 DR MGD; MGI:1919760; 2610528O17R1k.
 DR InterPro; IPR000508; SigPase.
 DR InterPro; IPR000223; SigPase_S26A.
 DR Pfam; PF00461; Peptidase_S26; 1.
 DR PRINTS; PR00727; LEADERPTASE.
 DR PROSITE; PS00761; SPASE I 3; 1.
 SQ SEQUENCE 166 AA; 18502 MW; 5627EB1F1BD80E3D CRC64;

Query Match 97.1%; Score 33; DB 11; Length 166;
 Best Local Similarity 80.0%; Pred. No. 24;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 Db 21 CIAHC 25

RESULT 11
 P91335 PRELIMINARY; PRT; 88 AA.
 ID P91335
 AC P91335
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 9.8 kDa protein.
 GN F55C7.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Du Z., Le T.;
 RT "The sequence of C. elegans cosmid F55C7";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U80436; AAC71107.1; -;
 SQ SEQUENCE 88 AA; 9771 MW; 59FB61FDFD7C0CE3 CRC64;

Query Match 91.2%; Score 31; DB 5; Length 88;
 Best Local Similarity 80.0%; Pred. No. 35;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Mon Jun 23 10:09:21 2003

Q9FUG3;
 AC 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Phytochelatin synthase (Fragment).
 OS Typha latifolia (Common cattail).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Typhaceae; Typha.
 OX NCBI_TaxID=4733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y.-W., Wong Y.-S., Tam F.-Y.;
 RT "Cloning of PC synthase Gene from Typha latifolia.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF308658; AAG22095.1; -;
 FT NON_TER 1
 FT NON_TER 133 133
 SQ SEQUENCE 133 AA; 14846 MW; 81712CC652EA0E54 CRC64;

Query Match 91.2%; Score 31; DB 10; Length 133;
 Best Local Similarity 80.0%; Pred. No. 49;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 Db 64 CIAHC 68

RESULT 13
 Q8QRA1 PRELIMINARY; PRT; 138 AA.
 ID Q8QRA1
 AC Q8QRA1
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE S1 glycoprotein (Fragment).
 OS Avian infectious bronchitis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11120;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA/3760/99;
 RA Schikora B.M., Shih L.M., Hietala S.K.;
 RT "Genetic Diversity of Avian Infectious Bronchitis Virus California
 RT Variants Isolated between 1988 and 2001 based on the S1 subunit of the
 RT spike glycoprotein.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF488289; AAM09489.1; -;
 FT NON_TER 1
 FT NON_TER 138 138
 SQ SEQUENCE 138 AA; 15675 MW; DBB14E0B65E37895 CRC64;

Query Match 91.2%; Score 31; DB 12; Length 138;
 Best Local Similarity 80.0%; Pred. No. 50;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
 Db 51 CIAHC 55

RESULT 14
 Q9JJ96 PRELIMINARY; PRT; 141 AA.
 ID Q9JJ96
 AC Q9JJ96;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Brain cDNA, clone MNCB-0671.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL;
RA Osada N., Kuseida J., Tanuma R., Ito A., Hirata M., Sugano S.;
RA Hashimoto K.; full-length cDNA clones from mouse brain cDNA library
RT made by oligo-capping method.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041649; BAA95101.1; -
DR InterPro; IPR001526; LV6_UPAR.
DR SMART; SM00134; LU; 1
SQ SEQUENCE 141 AA; 15331 MW; C8716DDC652288ED CRC64;

Query Match 91.2%; Score 31; DB 11; Length 141;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 137 CLAHC 141

RESULT 15
Q8QRAO PRELIMINARY; PRT; 141 AA.
AC Q8QRAO;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE S1 glycoprotein (fragment).
OS Avian infectious bronchitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11120;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CA/3761/99;
RA Schikora B.M., Shih L.M., Hietala S.K.;
RT "Genetic Diversity of Avian Infectious Bronchitis Virus California Variants isolated between 1988 and 2001 based on the S1 subunit of the spike glycoprotein.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF488290; AAM09490.1; -
FT NON TER 1
FT NON TER 141
SQ SEQUENCE 141 AA; 15043 MW; EBC400EB34982E22 CRC64;

Query Match 91.2%; Score 31; DB 12; Length 141;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 54 CVAHC 58

RESULT 16
O44981 PRELIMINARY; PRT; 173 AA.
AC O44981;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 20.2 kDa protein.
GN C54G6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
```

```

RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Wilson R., Tin-Wollam A., Graves T., Ozeraky P.;
RT "The sequence of C. elegans cosmid C54G6.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043698; AAB97561.1; -
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 20156 MW; 053B67461C2D59F3 CRC64;

Query Match 91.2%; Score 31; DB 5; Length 173;
Best Local Similarity 80.0%; Pred. No. 61;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 101 CVAHC 105

RESULT 17
O66549 PRELIMINARY; PRT; 182 AA.
AC O66549;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein AQ_159.
GN AQ_159.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.";
RL Nature 392:353-358 (1998).
DR EMBL; AE000676; AAC06513.1; -
DR InterPro; IPR004136; Znpnop_dioxygen.
DR InterPro; IPR003009; FMN_enzyme.
DR Pfam; PF03060; NPD; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 182 AA; 20422 MW; 8807CAA59CCE78C8 CRC64;

Query Match 91.2%; Score 31; DB 16; Length 182;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVAHC 5
Db 98 CVAHC 102

RESULT 18
Q9USD6 PRELIMINARY; PRT; 199 AA.
ID Q9USD6
AC Q9USD6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```



```

Query Match          91.2%; Score 31; DB 16; Length 400;
Best Local Similarity 80.0%; Pred.No. 1.2e+02;
Matches      4; Conservative    1; Mismatches     0; Indels   0; Gaps   0;

QY      1 CVAHC 5
DB      248 CVSHC 252
        |||::||
        |||::||

RESULT 23
Q20629 PRELIMINARY; PRT; 414 AA.
ID Q20629 AC Q20629;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-WAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical 47.8 kDa protein.
OS F49E8.2.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiidae; Pelodierinae; Caenorhabditis.
NCBI_TaxID=6239;
[1]_RN RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]_RN RN [2]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Murray J., Wohlmann P.;
RT "The sequence of C. elegans cosmid F49E8.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
[3]_RN RN [3]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; U61949; AAB03150.2; ...
DR InterPro; IPR002109; Glutaredoxin.
DR InterPro; IPR000822; Znf C2H2.
DR PROSITE; PS00195; GLUTAREDOXIN; UNKNOWN 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ Hypothetical protein.
KW SEQUENCE 414 AA; 47795 MW; AABCACG3C9A21EA8C CRC64;

Query Match          91.2%; Score 31; DB 5; Length 414;
Best Local Similarity 80.0%; Pred.No. 1.3e+02;
Matches      4; Conservative    1; Mismatches     0; Indels   0; Gaps   0;

QY      1 CVAHC 5
DB      6 CVSHC 10
        |||::||
        |||::||

RESULT 24
Q9FI40 PRELIMINARY; PRT; 443 AA.
ID Q9FI40 AC Q9FI40;
DT 01-WAR-2001 (TREMBLrel. 16, Created)
DT 01-WAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Acetyl-CoA:benzylalcohol acetyltransferase-like-protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
NCBI_TaxID=3702;
[1]_RN RN [1]_SEQUENCE FROM N.A.

```

```
Db          107 CLARC 111
          |:|
Search completed: June 20, 2003, 20:57:06
Job time : 28.3636 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 38.3182 Seconds
(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-20
Perfect score: 40
Sequence: 1 CHAVDC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : A Geneseq 101002.*

1:	/SID32/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2:	/SID32/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3:	/SID32/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4:	/SID32/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5:	/SID32/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6:	/SID32/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7:	/SID32/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8:	/SID32/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9:	/SID32/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10:	/SID32/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11:	/SID32/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12:	/SID32/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13:	/SID32/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14:	/SID32/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15:	/SID32/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16:	/SID32/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17:	/SID32/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18:	/SID32/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19:	/SID32/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20:	/SID32/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21:	/SID32/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22:	/SID32/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23:	/SID32/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	6	20	AA17148
2	40	100.0	6	21	AA173791
3	40	100.0	6	22	AA173791
4	40	100.0	6	23	AA173791
5	36	90.0	6	22	AA173791
6	36	90.0	6	22	AA173791
7	36	90.0	95	22	AA173791
8	36	90.0	217	22	AA173791
9	35	87.5	1798	19	AA173791
10	35	87.5	1798	23	AA173791

11	35	87.5	1798	23	AA173791
12	34	85.0	6	19	AA173791
13	34	85.0	6	20	AA173791
14	34	85.0	6	21	AA173791
15	34	85.0	6	22	AA173791
16	34	85.0	6	23	AA173791
17	34	85.0	78	22	AA173791
18	34	85.0	649	22	AA173791
19	33	82.5	6	22	AA173791
20	33	82.5	6	23	AA173791
21	33	82.5	85	22	AA173791
22	33	82.5	142	21	AA173791
23	33	82.5	142	22	AA173791
24	33	82.5	142	22	AA173791
25	33	82.5	142	22	AA173791
26	33	82.5	211	23	AA173791
27	33	82.5	334	22	AA173791
28	33	82.5	715	22	AA173791
29	33	82.5	823	21	AA173791
30	33	82.5	866	22	AA173791
31	33	82.5	933	22	AA173791
32	33	82.5	958	22	AA173791
33	33	82.5	965	21	AA173791
34	33	82.5	977	18	AA173791
35	33	82.5	1188	21	AA173791
36	32	80.0	226	12	AA173791
37	32	80.0	238	22	AA173791
38	32	80.0	246	12	AA173791
39	32	80.0	246	19	AA173791
40	32	80.0	246	20	AA173791
41	32	80.0	246	20	AA173791
42	32	80.0	246	20	AA173791
43	32	80.0	246	22	AA173791
44	32	80.0	454	23	AA173791
45	31	77.5	690	5	AA173791
46	31	77.5	6	20	AA173791
47	31	77.5	6	20	AA173791
48	31	77.5	6	21	AA173791
49	31	77.5	6	22	AA173791
50	31	77.5	6	23	AA173791
51	31	77.5	6	23	AA173791
52	31	77.5	7	19	AA173791
53	31	77.5	7	20	AA173791
54	31	77.5	7	21	AA173791
55	31	77.5	7	22	AA173791
56	31	77.5	7	22	AA173791
57	31	77.5	7	23	AA173791
58	31	77.5	7	23	AA173791
59	31	77.5	8	20	AA173791
60	31	77.5	8	21	AA173791
61	31	77.5	8	22	AA173791
62	31	77.5	8	22	AA173791
63	31	77.5	8	23	AA173791
64	31	77.5	8	23	AA173791
65	31	77.5	9	20	AA173791
66	31	77.5	9	21	AA173791
67	31	77.5	9	22	AA173791
68	31	77.5	9	22	AA173791
69	31	77.5	9	23	AA173791
70	31	77.5	9	23	AA173791
71	31	77.5	38	22	AA173791
72	31	77.5	45	22	AA173791
73	31	77.5	63	22	AA173791
74	31	77.5	81	22	AA173791
75	31	77.5	100	22	AA173791

ALIGNMENTS

RESULT 1
AA17148
ID AA17148 standard; peptide; 6 AA.

XX AC AAY17148;
 XX DT 20-SEP-1999 (first entry)
 XX DE Cadherin-mediated adhesion modulating cyclic peptide.
 XX DE Cell adhesion modulation; CAM; synaptic stability; cadherin;
 KW cadherin-mediated adhesion; drug delivery; cell adhesion; tumour;
 KW wound healing; neurite outgrowth; cyclic; circular.
 XX OS Synthetic.
 XX PN W09933875-A1.
 XX PD 08-JUL-1999.
 XX PF 23-DEC-1998; 98WO-CA01207.
 XX PR 23-DEC-1997; 97US-0996679.
 XX PA (UYMC-) UNIV MCGILL.
 XX PI Blaschuk OW, Gour BJ;
 XX PF WPI; 1999-430231/36.
 XX DR Cyclic peptide cell adhesion modulating agents, useful for
 PT modulating synaptic stability
 XX PS Claim 11; Page 64; 144pp; English.
 XX CC The invention provides cyclic peptide cell adhesion modulating (CAM)
 CC agents that comprises a His-Ala-Val recognition sequence. Also provided
 CC is a method for inhibiting synaptic stability in a mammal that comprises
 CC administering to a mammal a therapeutically effective amount of a CAM
 CC agent that inhibits cadherin-mediated adhesion, where the agent comprises
 CC a cyclic peptide having a peptide ring, and where the sequence His-Ala-
 CC Val is present within the peptide ring. The cyclic peptides are cell
 CC adhesion modulating agents that inhibit cadherin-mediated adhesion. They
 CC can be used in a method for inhibiting synaptic stability in mammals. The
 CC agents can be used to treat diseases or other conditions characterized by
 CC undesirable cell adhesion or to facilitate drug delivery to a specific
 CC tissue or tumour. Alternatively the agents may be used to enhance cell
 CC adhesion (e.g. to supplement or replace sutures or to facilitate wound
 CC healing) or to enhance or direct neurite outgrowth.
 XX SQ Sequence 6 AA;

Query Match 100.0%; Score 40; DB 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 |||||
 Db 1 CHAVDC 6

RESULT 2
 AAY73791
 ID AAY73791 standard; Peptide; 6 AA.
 XX AC AAY73791;
 XX DT 07-MAR-2000 (first entry)
 XX DE N-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:297.
 XX DE Diagnosis; cancer; OB-cadherin; N-cadherin; E-cadherin; CAR; cyclic;
 KW cell adhesion recognition sequence; leukaemia; metastasis;
 KW prostate cancer; breast cancer; ovarian cancer; carcinoma.
 XX OS Synthetic.

OS Homo sapiens.
 XX WO9957565-A2.
 XX PD 11-NOV-1999.
 XX PF 05-MAY-1999; 99WO-CA00362.
 XX PR 05-MAY-1998; 98US-0073040.
 XX PR 06-NOV-1998; 98US-0187859.
 XX PR 20-JAN-1999; 99US-0234395.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Blaschuk OW, Gour BJ, Byers S;
 XX PF WPI; 2000-062165/05.
 XX DR Detecting expression of OB-cadherin or N-cadherin used for diagnosing
 PT and evaluating cancer
 XX PS Disclosure; Page 21; 56pp; English.
 XX CC The present invention describes methods which have been developed for
 CC detecting and evaluating cancer that are based on the finding that
 CC OB-cadherin and N-cadherin are expressed by metastatic carcinoma cells,
 CC but not by highly differentiated, poorly invasive carcinomas. A method
 CC for determining the presence or absence of a cancer in a patient
 CC comprises: (1) contacting a biological sample from the patient with a
 CC binding agent that specifically binds to OB- or N-cadherin OR with an
 CC oligonucleotide that hybridizes to a polynucleotide that encodes OB- or
 CC N-cadherin; and (2) detecting in the sample an amount of polypeptide
 CC that binds to the binding agent OR the amount of polynucleotide that
 CC hybridizes to the oligonucleotide, relative to a predetermined cut-off
 CC value, and determining the presence or absence of cancer in the patient
 CC from this. The methods from the present invention can be used to
 CC determine the metastatic potential of a cancer. The methods may be used
 CC to detect a metastatic cancer in a patient, to monitor progression of a
 CC cancer, or to evaluate the metastatic potential of a cancer. Cancers
 CC which may be evaluated using the methods include leukemia, prostate
 CC cancer, breast cancer and ovarian cancer. AAY73501 to AAY73813 represent
 CC peptide sequences used in the exemplification of the present invention.
 CC Specifically, AAY73503 to AAY73789 represent OB-cadherin cell adhesion
 CC recognition (CAR) peptides and AAY73790 to AAY73808 represent N-cadherin
 CC CAR peptides. AAY7396 to AAY74915 represent PCR primers used in examples
 CC from the present invention.
 XX SQ Sequence 6 AA;

Query Match 100.0%; Score 40; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 |||||
 Db 1 CHAVDC 6

RESULT 3
 AAG65380
 ID AAG65380 standard; peptide; 6 AA.
 XX AC AAG65380;
 XX DT 30-NOV-2001 (first entry)
 XX DE Cyclic peptide with cadherin CAR sequence.
 XX DE Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 KW cytostatic; vulnerable; immunomodulator; vasotropic; neuroprotective;
 KW cerebroprotective; muscular; cyclic.
 XX OS Synthetic.

XX Key Location/Qualifiers
 PH Modified-site 1 /note= "putative N-terminal acetylation or
 FT alkoxybenzylation"
 FT Modified-site 6 /note= "putative C-terminal amide or ester"
 FT
 XX WO200153331-A2.
 XX 26-JUL-2001.
 XX 24-JAN-2001; 2001WO-US02508.
 XX 24-JAN-2000; 2000US-0491078.
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 PI Hu Z;
 XX WPI; 2001-549899/61.
 XX Cell adhesion modulating agent used for enhancing delivery of drug to
 PT tumor comprises imidazole compounds -
 PT
 XX Claim 16; Page 146; 436pp; English.
 XX The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cyclic peptide with
 CC cadherin cell adhesion recognition (CAR) sequence.
 XX Sequence 6 AA;
 SQ Query Match 100.0%; Score 40; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CHAVDC 6
 Db |||||
 1 CHAVDC 6
 RESULT 4
 AAM47487
 ID AAM47487 standard; peptide; 6 AA.
 XX AAM47487;
 XX 12-FEB-2002 (first entry)
 XX Cyclic peptide endothelial cell adhesion modulator #5.
 DE Cadherin; cytostatic; gynecological; endometriosis;
 KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 KW cyclic.
 XX Synthetic.
 OS Key Location/Qualifiers
 PH Modified-site 1 /note= "Linked to residue 6 to form a cyclic peptide,"
 FT

FT N-terminal acetyl or N-terminal CH3-SO2 group"
 FT Modified-site 6 /note= "Linked to residue 1 to form a cyclic peptide,"
 FT C-terminal amide"
 XX WO200171146-A2.
 XX 18-OCT-2001.
 XX 09-APR-2001; 2001WO-US11669.
 XX 07-APR-2000; 2000US-0544782.
 XX (UYMC-) UNIV MCGILL.
 XX Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 PI WPI; 2002-049129/06.
 XX Modulating endothelial cell adhesion for inhibiting development of
 PT endometriosis, increasing blood flow to tumor in a mammal, by
 PT contacting cell with a cyclic peptide having cadherin cell adhesion
 PT recognition sequence -
 PT
 XX Claim 6; Page 64; 139pp; English.
 XX The present invention relates to a method for modulating endothelial cell
 CC adhesion. The method comprises contacting an endothelial cell with a
 CC modulating peptide comprising a cadherin cell adhesion recognition
 CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 CC for inhibiting angiogenesis, increasing vasopermeability, increasing the
 CC blood flow to a tumour, disrupting neovasculation and inhibiting the
 CC development of endometriosis in a mammal. The modulating peptide reduces
 CC unwanted endothelial adhesion occurring between tumour cells, tumour
 CC cells and normal cells, normal cells as a result of surgery, injury,
 CC chemotherapy, disease and inflammation. The present sequence is one such
 CC modulating peptide.
 XX Sequence 6 AA;
 SQ Query Match 100.0%; Score 40; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CHAVDC 6
 Db |||||
 1 CHAVDC 6
 RESULT 5
 AAG65381
 ID AAG65381 standard; peptide; 6 AA.
 XX AAG65381;
 XX 30-NOV-2001 (first entry)
 XX Cyclic control peptide sequence.
 DE Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 KW cytostatic; vulnary; immunomodulator; vasotropic; neuroprotective;
 KW cerebroprotective; muscular; cyclic.
 XX Synthetic.
 OS Key Location/Qualifiers
 PH Modified-site 1 /note= "putative N-terminal acetylation or
 FT alkoxybenzylation"
 FT Modified-site 6 /note= "putative C-terminal amide or ester"
 FT
 XX WO200153331-A2.

XX PD 26-JUL-2001.
 XX PF 24-JAN-2001; 2001WO-US02508.
 XX PR 24-JAN-2000; 2000US-0491078.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 XX PI Hu Z;
 XX PR WPI; 2001-549899/61.
 XX DR
 XX PT Cell adhesion modulating agent used for enhancing delivery of drug to
 PT tumor comprises imidazole compounds -
 XX Examples; Fig 3A-I; 436pp; English.
 XX The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cycliccontrol peptide.
 XX SQ Sequence 6 AA;
 Query Match 90.0%; Score 36; DB 22; Length 6;
 Best Local Similarity 83.3%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CHAVDC 6
 DB 1 CHGVDC 6
 RESULT 6
 AAM47488
 ID AAM47488 standard; peptide; 6 AA.
 AC AAM47488;
 XX DT 12-FEB-2002 (first entry)
 XX DE Cyclic control peptide #7.
 XX KW Cadherin; cytostatic; gynecological; endometriosis;
 KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 KW cyclic.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Linked to residue 6 to form a cyclic peptide,
 FT N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Linked to residue 1 to form a cyclic peptide,
 FT C-terminal amide".
 XX WO200177146-A2.
 XX PN 18-OCT-2001.
 XX PD 09-APR-2001; 2001WO-US11669.
 XX PF

XX PR 07-APR-2000; 2000US-0544782.
 XX PA (UYMC-) UNIV MCGILL.
 XX PI Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 XX PR WPI; 2002-049129/06.
 XX DR
 XX PT Modulating endothelial cell adhesion for inhibiting development of
 PT endometriosis, increasing blood flow to tumor in a mammal, by
 PT contacting cell with a cyclic peptide having cadherin cell adhesion
 PT recognition sequence -
 XX PS Disclosure; Fig 3; 139pp; English.
 XX The present invention relates to a method for modulating endothelial cell
 CC adhesion. The method comprises contacting an endothelial cell with a
 CC modulating peptide comprising a cadherin cell adhesion recognition
 CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 CC for inhibiting angiogenesis, increasing vasopermeability, increasing
 CC blood flow to a tumour, disrupting neovasculture and inhibiting the
 CC development of endometriosis in a mammal. The modulating peptide reduces
 CC unwanted endothelial adhesion occurring between tumour cells, tumour
 CC cells and normal cells, normal cells as a result of surgery, injury,
 CC chemotherapy, disease and inflammation. The present sequence is a peptide
 CC used as a control for illustrating the present invention.
 XX SQ Sequence 6 AA;
 Query Match 90.0%; Score 36; DB 23; Length 6;
 Best Local Similarity 83.3%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CHAVDC 6
 DB 1 CHGVDC 6
 RESULT 7
 AAM91705
 ID AAM91705 standard; Protein; 95 AA.
 AC AAM91705;
 XX DT 07-NOV-2001 (first entry)
 XX DE Human immune/haematopoietic antigen SEQ ID NO:19298.
 XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 XX OS Homo sapiens.
 XX WO200157182-A2.
 XX PN 09-AUG-2001.
 XX PD 17-JAN-2001; 2001WO-US01354.
 XX PF 31-JAN-2000; 2000US-0179065.
 XX PR 04-FEB-2000; 2000US-0180628.
 XX PR 24-FEB-2000; 2000US-0184684.
 XX PR 02-MAR-2000; 2000US-0186350.
 XX PR 16-MAR-2000; 2000US-0189874.
 XX PR 17-MAR-2000; 2000US-0190076.
 XX PR 18-APR-2000; 2000US-0198123.
 XX PR 19-MAY-2000; 2000US-0205515.
 XX PR 07-JUN-2000; 2000US-0209467.
 XX PR 28-JUN-2000; 2000US-0214886.
 XX PR 30-JUN-2000; 2000US-0215135.
 XX PR 07-JUL-2000; 2000US-0216647.
 XX PR 07-JUL-2000; 2000US-0216880.

PR	11-JUL-2000;	2000US-02117487.	PR	20-OCT-2000;	2000US-0241826.
PR	11-JUL-2000;	2000US-02117487.	PR	01-NOV-2000;	2000US-0244617.
PR	14-JUL-2000;	2000US-02117496.	PR	08-NOV-2000;	2000US-0246474.
PR	14-JUL-2000;	2000US-02118290.	PR	08-NOV-2000;	2000US-0246475.
PR	26-JUL-2000;	2000US-0220963.	PR	08-NOV-2000;	2000US-0246476.
PR	26-JUL-2000;	2000US-0220964.	PR	08-NOV-2000;	2000US-0246477.
PR	14-AUG-2000;	2000US-0224518.	PR	08-NOV-2000;	2000US-0246478.
PR	14-AUG-2000;	2000US-0224519.	PR	08-NOV-2000;	2000US-0246523.
PR	14-AUG-2000;	2000US-0225213.	PR	08-NOV-2000;	2000US-0246524.
PR	14-AUG-2000;	2000US-0225214.	PR	08-NOV-2000;	2000US-0246525.
PR	14-AUG-2000;	2000US-0225266.	PR	08-NOV-2000;	2000US-0246527.
PR	14-AUG-2000;	2000US-0225267.	PR	08-NOV-2000;	2000US-0246528.
PR	14-AUG-2000;	2000US-0225268.	PR	08-NOV-2000;	2000US-0246532.
PR	14-AUG-2000;	2000US-0225270.	PR	08-NOV-2000;	2000US-0246532.
PR	14-AUG-2000;	2000US-0225447.	PR	08-NOV-2000;	2000US-0246609.
PR	14-AUG-2000;	2000US-0225758.	PR	08-NOV-2000;	2000US-0246610.
PR	14-AUG-2000;	2000US-0225759.	PR	08-NOV-2000;	2000US-0246611.
PR	18-AUG-2000;	2000US-0226279.	PR	08-NOV-2000;	2000US-0246613.
PR	22-AUG-2000;	2000US-0226681.	PR	17-NOV-2000;	2000US-0249207.
PR	22-AUG-2000;	2000US-0226688.	PR	17-NOV-2000;	2000US-0249208.
PR	22-AUG-2000;	2000US-0226688.	PR	17-NOV-2000;	2000US-0249209.
PR	22-AUG-2000;	2000US-0227182.	PR	17-NOV-2000;	2000US-0249210.
PR	23-AUG-2000;	2000US-0227009.	PR	17-NOV-2000;	2000US-0249211.
PR	30-AUG-2000;	2000US-0228924.	PR	17-NOV-2000;	2000US-0249212.
PR	01-SEP-2000;	2000US-0229287.	PR	17-NOV-2000;	2000US-0249213.
PR	01-SEP-2000;	2000US-0229343.	PR	17-NOV-2000;	2000US-0249214.
PR	01-SEP-2000;	2000US-0229344.	PR	17-NOV-2000;	2000US-0249215.
PR	01-SEP-2000;	2000US-0229345.	PR	17-NOV-2000;	2000US-0249216.
PR	05-SEP-2000;	2000US-0229509.	PR	17-NOV-2000;	2000US-0249217.
PR	05-SEP-2000;	2000US-0229513.	PR	17-NOV-2000;	2000US-0249218.
PR	06-SEP-2000;	2000US-0230437.	PR	17-NOV-2000;	2000US-0249218.
PR	06-SEP-2000;	2000US-0230438.	PR	17-NOV-2000;	2000US-0249244.
PR	08-SEP-2000;	2000US-0231242.	PR	17-NOV-2000;	2000US-0249245.
PR	08-SEP-2000;	2000US-0231243.	PR	17-NOV-2000;	2000US-0249264.
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	2000US-0231413.	PR	17-NOV-2000;	2000US-0249297.
PR	08-SEP-2000;	2000US-0231414.	PR	17-NOV-2000;	2000US-0249299.
PR	08-SEP-2000;	2000US-0232080.	PR	17-NOV-2000;	2000US-0249300.
PR	08-SEP-2000;	2000US-0232081.	PR	01-DEC-2000;	2000US-0250160.
PR	12-SEP-2000;	2000US-0231968.	PR	01-DEC-2000;	2000US-0250191.
PR	14-SEP-2000;	2000US-0232397.	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	2000US-0232398.	PR	05-DEC-2000;	2000US-0251988.
PR	14-SEP-2000;	2000US-0232399.	PR	05-DEC-2000;	2000US-0256719.
PR	14-SEP-2000;	2000US-0232400.	PR	06-DEC-2000;	2000US-0251479.
PR	14-SEP-2000;	2000US-0232401.	PR	08-DEC-2000;	2000US-0251856.
PR	14-SEP-2000;	2000US-0233063.	PR	08-DEC-2000;	2000US-0251868.
PR	14-SEP-2000;	2000US-0233064.	PR	08-DEC-2000;	2000US-0251869.
PR	14-SEP-2000;	2000US-0233065.	PR	08-DEC-2000;	2000US-0251989.
PR	21-SEP-2000;	2000US-0234223.	PR	08-DEC-2000;	2000US-0251990.
PR	21-SEP-2000;	2000US-0234224.	PR	11-DEC-2000;	2000US-0254097.

CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (i) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.

CC SQ Sequence 95 AA;
Query Match 90.0%; Score 36; DB 22; Length 95;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 42 CHAVEC 47

RESULT 8
AAAM39069
ID AAM39069 standard; Protein; 217 AA.

XX AC AAM39069;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 2214.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX OS Homo sapiens.
XX PN WO200153312-A1.

XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang JT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX DR N-PSDB; AAI58225.

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Example 4; SEQ ID NO 2214; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

CC SQ Sequence 217 AA;
Query Match 90.0%; Score 36; DB 22; Length 217;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 164 CHAVEC 169

RESULT 9
AAW50896
ID AAW50896 standard; Protein; 1798 AA.

XX AC AAW50896;
XX DT 07-DEC-1998 (first entry)
XX DE Human laminin B2 chain.

XX Laminin; human; beta-amyloid; amyloidosis; Alzheimer's disease;
KW Down's syndrome; hereditary cerebral haemorrhage; inflammation;
KW malignancy; Familial Mediterranean Fever; multiple myeloma;
KW type II diabetes; prion disease; Creutzfeldt-Jacob disease; CJD;
KW Gertsmann-Straussler syndrome; Kuru; scrapie; haemodialysis;
KW carpal tunnel syndrome; senile cardiac amyloid polynuropathy;
KW Familial Amyloidotic Polynuropathy; thyroid carcinoma; diagnosis;
KW therapy.

XX OS Homo sapiens.

XX PN WO9815179-A1.

XX PD 16-APR-1998.

XX PF 08-OCT-1997; 97WO-US18145.

XX PR 08-OCT-1996; 96US-0027981.

XX (UNIW) UNIV WASHINGTON.

XX Castillo G, Snow AD;

XX WPI; 1998-240534/21.

XX Use of laminin and fragments - for developing products for use in
PT the diagnosis and treatment of amyloid disease, e.g. Alzheimer's
PT disease or CJD
XX Claim 15; Page 98-101; 132pp; English.

XX This is the amino acid sequence of the human laminin B2 chain. The
CC primary object of the invention is to use laminin, laminin-derived
CC protein fragments and/or laminin-derived polypeptides as potent
CC inhibitors of amyloid formation, deposition, accumulation and/or
CC persistence in Alzheimer's disease and other amyloidoses. The
CC laminin products (see AAW50888-98) may include mouse or human laminin
CC A or A1 chain, laminin B1 or B2 chain, laminin A2 chain (merosin),
CC laminin G1 chain, the globular repeats of the laminin A1 chain and
CC the beta-amyloid binding domain of the laminin A chain. A claimed

method for treating an amyloid disease comprises administering a polypeptide having a conformational similarity to a fragment of a laminin protein. A method for diagnosing an amyloid disease involves determining levels of laminin in a sample. Production of laminin or its fourth globular repeat in vivo provides a method for in vivo inhibition of beta-amyloid amyloidosis. The products and methods can be used for the diagnosis, prognosis, monitoring and treatment of amyloidoses such as Alzheimer's disease, Down's syndrome and hereditary cerebral haemorrhage with amyloidosis of the Dutch type (where the specific amyloid is the beta-amyloid protein), the amyloidosis associated with chronic inflammation, various forms of malignancy and Familial Mediterranean Fever (AA amyloid or inflammation-associated amyloidosis), the amyloidosis associated with multiple myeloma and other B-cell abnormalities (AL amyloid), the amyloidosis associated with type II diabetes (amylin or islet amyloid), the amyloidosis associated with prion diseases including Creutzfeldt-Jacob disease, Gerstmann-Strausler syndrome, kuru and animal scrapie (PrP amyloid), the amyloidosis associated with long-term haemodialysis and carpal tunnel syndrome (beta 2-microglobulin amyloid), the amyloidosis associated with senile cardiac amyloid and Familial Amyloidotic Polyneuropathy (prealbumin or transthyretin amyloid), and the amyloidosis associated with endocrine tumours such as medullary carcinoma of the thyroid (variant of procalcitonin).

SQ Sequence 1798 AA;

Query Match 87.5%; Score 35; DB 19; Length 1798;
Best Local Similarity 83.3%; Pred. NO. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
||| |||

Db 1140 CHACDC 1145

RESULT 10
AAU84346
ID AAU84346 standard; Protein; 1798 AA.
XX AC AAU84346;
XX DT 08-MAY-2002 (first entry)
XX DE Protein LAMB2 differentially expressed in breast cancer tissue.
XX KW Human; diagnosis of breast cancer; endometrial cancer; breast tumour;
XX KW MAI; mitotic activity index; cytostatic.
XX OS Homo sapiens.
XX PN WO200210436-A2.
XX PD 07-FEB-2002.
XX PF 27-JUL-2001; 2001WO-US23642.
XX PR 28-JUL-2000; 2000US-222093P.
XX PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX PA (BAK/) BAK J.
XX FI Baak J, Mutter GL;
XX DR WPI; 2002-180084/23.
XX DR N-PSDB; ABK35566.
XX PT Diagnosing breast cancer comprises determining expression of nucleic acid molecules or expression products that are differentially expressed in normal and malignant tissue -
XX PS Claim 37; Page 177-183; 219pp; English.

The present invention relates to a method for diagnosing breast cancer in a subject suspected of having endometrial cancer. The method comprises determining the expression of a set of human genes or expression products in an endometrial sample suspected of being cancerous. The human genes of the invention are differentially expressed in breast tumours characterised as high or low MAI (mitotic activity index). These sets of genes can be used to discriminate between high and low MAI breast tumours. The invention also provides DNA and protein microarrays for analysing the expression of the human genes and their protein products. The methods and arrays are useful for the diagnosis and prognosis of endometrial cancer, selecting and monitoring treatment regimes, and identification of compounds useful for the treatment of endometrial cancer. AAU84311-AAU84361 represent the human proteins of the invention that are differentially expressed in breast cancer tissue.

SQ Sequence 1798 AA;

Query Match 87.5%; Score 35; DB 23; Length 1798;
Best Local Similarity 83.3%; Pred. NO. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
||| |||

Db 1140 CHACDC 1145

RESULT 11
AAM50360
ID AAM50360 standard; Protein; 1798 AA.
XX AC AAM50360;
XX DT 18-FEB-2002 (first entry)
XX DE Human laminin-15 beta 2 chain.
XX KW Laminin-15; human; retina; eye; therapy; ophthalmological;
XX KW antiinflammatory; rod dystrophy; rod-cone dystrophy;
XX KW retinitis pigmentosa; macular degeneration; retinal detachment.
XX OS Homo sapiens.
XX PN WO200183516-A1.
XX PD 08-NOV-2001.
XX PF 01-MAY-2001; 2001WO-US13943.
XX PR 01-MAY-2000; 2000US-200863P.
XX PA (MASS-) MASSACHUSETTS GEN HOSPITAL.
XX PI Burgeson RE, Brunkan W, Champliand M, Hunter D;
XX DR WPI; 2002-041478/05.
XX DR N-PSDB; AAI70819.
XX PT Novel substantially pure preparation comprising laminin having laminin chain alpha 5, beta 2, and gamma 3, useful for treating retinal disorders such as retinitis pigmentosa, macular degeneration, retinal detachment -
XX PS Disclosure; Fig 4A; 58pp; English.
XX CC The present sequence is that of the beta 2 chain of human laminin-15, a novel member of the laminin family that is produced in the retina.
XX CC The retina produces 2 novel laminin trimers: laminin-14 (alpha 4, beta 2, gamma 3) and laminin-15 (alpha 5, beta 2, gamma 3). These are expressed within the inter-photoreceptor matrix and in the outer plexiform layer, and may serve to stabilise retinal synapses.
XX CC The invention provides laminin-15 preparations and cells comprising a nucleic acid encoding the laminin alpha 5, beta 2 and gamma 3

CC chains, and which are capable of producing laminin-15. The
 CC laminin-15 preparation is used in claimed methods of: increasing
 CC retinal immunophoreceptor matrix stability; increasing the
 CC stability of retinal photoreceptor compounds, especially an outer
 CC segment, inner segment or synapse; increasing retinal adhesion;
 CC treating a disorder associated with retinal degeneration,
 CC especially rod dystrophy, rod-cone dystrophy, retinitis pigmentosa,
 CC macular degeneration and retinal detachment; increasing the
 CC stability of synapses of the central nervous system or peripheral
 CC nervous system; stimulating neuroregeneration, axon outgrowth or
 CC synapse formation; preparing an implant, e.g. a catheter,
 CC artificial joint, retinal implant, timed releasing device, neural
 CC cell growth guide or artificial tissue, by coating with the
 CC laminin-15 preparation; and increasing photosensitivity by
 CC implanting a tip coated with the laminin-15 preparation into an
 CC eye. The laminin may be recombinant, and the 3 chains co-expressed
 CC in the same cell or expressed in different cells.

XX SQ Sequence 1798 AA;
 Query Match 87.5%; Score 35; DB 23; Length 1798;
 Best Local Similarity 83.3%; Pred. No. 6.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
 AAW23982
 ID AAW23982 standard; peptide; 6 AA.

XX AC AAW23982;

XX DT 23-JUL-1998 (first entry)

XX DE Cadherin-mediated cell adhesion modulating peptide 4.

XX KW Cadherin-mediated; cell adhesion; drug delivery; treatment; cancer;
 XX KW angiogenesis; skin grafting; neurite growth; veterinary medicine;
 XX KW wound healing; tumour; metastasis; carcinoma; leukaemia; modulation;
 XX KW organ transplant; neurological disease.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Disulfide-bond 1 /note= "disulphide bridge"
 FT Modified-site 6 /note= "C-terminal amide"
 FT Disulfide-bond 6 /note= "disulphide bridge"

XX WO9802452-A2.

XX PD 22-JAN-1998.

XX PF 11-JUL-1997; 97WO-CA00489.

XX PR 12-JUL-1996; 96US-0021612.

XX PA (UYMC-) UNIV MCGILL.

XX PI Blaschuk OW, Gour BJ;

XX DR WPI; 1998-110522/10.

XX PT Cyclic peptide(s) that modulate cadherin-mediated cell adhesion -
 XX PT used to improve drug delivery through skin, to the brain etc., for
 XX PT treatment of cancer, angiogenesis etc., and to improve grafting of

PT foreign tissue or neurite growth
 XX Claim 11; Page: 97; 133pp; English.

XX This is a cyclised peptide capable of modulating cadherin-mediated cell
 CC adhesion. Cadherin-expressing cells can be detected from their reaction
 CC with the antibody that binds the peptides of the invention. These
 CC peptides are useful in human or veterinary medicine to modulate adhesion
 CC mediated by E-, N-, P- or R-cadherins (or other cadherins that include
 CC the amino acids HAV recognition sequence), specifically where these are
 CC expressed by epithelial, endothelial, neural or tumour cells or
 CC lymphocytes. The peptides which inhibit cell adhesion are used to improve
 CC delivery of drugs through the skin (such that the peptides enters the
 CC blood stream), to tumours (particularly ovarian or bladder tumours or
 CC melanoma) and to the brain. They are also used to treat cancer
 CC (carcinoma, leukaemia or melanoma), inhibit metastasis and also inhibit
 CC angiogenesis. The peptides that stimulate adhesion are used to improve
 CC wound healing to promote adherence of foreign tissues (skin grafts or
 CC organ transplants), and to improve adherence to tissue culture surfaces
 CC and bioreactors. Other uses of the peptides are to induce apoptosis in
 CC cadherin-expressing cells, increase or decrease neurite outgrowth, to
 CC treat spinal cord injuries and de-myelinating neurological diseases
 CC (specifically multiple sclerosis). They can be used for modulating the
 CC immune system (e.g. in cases of diabetes or rheumatoid arthritis), for
 CC preventing pregnancy and to increase vaso-permeability. The antibodies
 CC which bind to these peptides are also used to modulate cell adhesion and
 CC when coupled to a drug, to target the drug to cadherin-expressing cells.
 CC Diagnostic agents may also be delivered using the peptides in a skin
 CC patch, particularly where the agent generates a colour in contact with,
 CC e.g. cocaine, human immunodeficiency viral proteins, glucose or
 CC prostate-specific antigen, particularly for home-testing kits. The
 CC peptides make it possible to deliver a wide range of drugs through the
 CC skin, avoiding parenteral administration, by-passing the gastro-
 CC intestinal system and improving patient compliance.

XX SQ Sequence 6 AA;

Query Match 85.0%; Score 34; DB 19; Length 6;
 Best Local Similarity 83.3%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 |||||
 DB 1 CHAVSC 6;

RESULT 13

AAV17114

ID AAV17114 standard; peptide; 6 AA.

XX AC AAV17114;

XX DT 20-SEP-1999 (first entry)

XX DE Cadherin-mediated adhesion modulating cyclic peptide.

XX KW Cell adhesion modulation; CAM; synaptic stability; cadherin;
 KW cadherin-mediated adhesion; drug delivery; cell adhesion; tumour;
 KW wound healing; neurite outgrowth; cyclic; circular.

XX OS Synthetic.

XX PN WO9933875-A1.

XX PD 08-JUL-1999.

XX PF 23-DEC-1998; 98WO-CA01207.

XX PR 23-DEC-1997; 97US-0996679.

XX PA (UYMC-) UNIV MCGILL.

XX PI Blaschuk OW, Gour BJ;

XX WPI; 1999-430231/36.

XX Cyclic peptide cell adhesion modulating agents, useful for
PT modulating synaptic stability

XX Claim 11; Page 64; 144pp; English.

XX The invention provides cyclic peptide cell adhesion modulating (CAM)
CC agents that comprises a His-Ala-Val recognition sequence. Also provided
CC is a method for inhibiting synaptic stability in a mammal that comprises
CC administering to a mammal a therapeutically effective amount of a CAM
CC agent that inhibits cadherin-mediated adhesion, where the agent comprises
CC a cyclic peptide having a peptide ring, and where the sequence His-Ala-
CC Val is present within the peptide ring. The cyclic peptides are cell
CC adhesion modulating agents that inhibit cadherin-mediated adhesion. They
CC can be used in a method for inhibiting synaptic stability in mammals. The
CC agents can be used to treat diseases or other conditions characterized by
CC undesirable cell adhesion or to facilitate drug delivery to a specific
CC tissue or tumour. Alternatively the agents may be used to enhance cell
CC adhesion (e.g. to supplement or replace stitches or to facilitate wound
CC healing) or to enhance or direct neurite outgrowth.

XX SQ Sequence 6 AA;

Query Match 85.0%; Score 34; DB 20; Length 6;
Best Local Similarity 83.3%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||||
Db 1 CHAVSC 6

RESULT 14

AA73799
ID AAY73799 standard; Peptide; 6 AA.

XX AC AAY73799;

XX 07-MAR-2000 (first entry)

XX N-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:305.

XX Diagnosis; cancer; OB-cadherin; N-cadherin; E-cadherin; CAR; cyclic;
KW cell adhesion recognition sequence; leukaemia; metastasis;
KW prostate cancer; breast cancer; ovarian cancer; carcinoma.

XX Synthetic.

OS Homo sapiens.

XX WO9957565-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00362.

XX 05-MAY-1998; 98US-0073040.

XX 06-NOV-1998; 98US-0187859.

XX 20-JAN-1999; 99US-0234395.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuck OW, Gour BJ, Byers S;

XX WPI; 2000-062165/05.

XX Detecting expression of OB-cadherin or N-cadherin used for diagnosing
PT and evaluating cancer

XX Disclosure; Page 22; 56pp; English.

XX The present invention describes methods which have been developed for

CC detecting and evaluating cancer that are based on the finding that
CC OB-cadherin and N-cadherin are expressed by metastatic carcinoma cells,
CC but not by highly differentiated, poorly invasive carcinomas. A method
CC for determining the presence or absence of a cancer in a patient
CC comprises: (1) contacting a biological sample from the patient with a
CC binding agent that specifically binds to OB- or N-cadherin OR with an
CC oligonucleotide that hybridizes to a polynucleotide that encodes OB- or
CC N-cadherin; and (2) detecting in the sample an amount of polypeptide
CC that binds to the binding agent OR the amount of polynucleotide that
CC hybridizes to the oligonucleotide, relative to a predetermined cut-off
CC value, and determining the presence or absence of cancer in the patient
CC from this. The methods from the present invention can be used to
CC determine the metastatic potential of a cancer. The methods may be used
CC to detect a metastatic cancer in a patient, to monitor progression of a
CC cancer, or to evaluate the metastatic potential of a cancer. Cancers
CC which may be evaluated using the methods include leukemia, prostate
CC cancer, breast cancer and ovarian cancer. AAY73501 to AAY73813 represent
CC peptide sequences used in the exemplification of the present invention.
CC Specifically, AAY73503 to AAY73789 represent OB-cadherin cell adhesion
CC recognition (CAR) peptides and AAY73790 to AAY73808 represent N-cadherin
CC CAR peptides. AAY73906 to AAY73915 represent PCR primers used in examples
CC from the present invention.

XX SQ Sequence 6 AA;

Query Match 85.0%; Score 34; DB 21; Length 6;
Best Local Similarity 83.3%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||||
Db 1 CHAVSC 6

RESULT 15

AAG65398
ID AAG65398 standard; peptide; 6 AA.

XX AC AAG65398;

XX 30-NOV-2001 (first entry)

XX Cyclic peptide with cadherin CAR sequence.

XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
KW cytostatic; vulnery; immunomodulator; vasotrophic; neuroprotective;
KW cerebroprotective; muscular; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "putative N-terminal acetylation or
FT alkoxybenzylolation"

FT Modified-site 6 /note= "putative C-terminal amide or ester"

XX WO200153331-A2.

XX 26-JUL-2001.

XX 24-JAN-2001; 2001WO-US02508.

XX 24-JAN-2000; 2000US-0491078.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Mithaud SD, Wang S;
PI Hu Z;

XX WPI; 2001-549899/61.

XX Cell adhesion modulating agent used for enhancing delivery of drug to

PT tumor comprises imidazole compounds -

PS Examples; Fig 3A-I; 436pp; English.

XX The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cyclic peptide with
 CC cadherin cell adhesion recognition (CAR) sequence.

SQ Sequence 6 AA;

Query Match 85.0%; Score 34; DB 22; Length 6;
 Best Local Similarity 83.3%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 |||||
 Db 1 CHAVSC 6

RESULT 16

AA47505
 ID AAM47505 standard; peptide; 6 AA.

AC AAM47505;

DT 12-FEB-2002 (first entry)

DE Cyclic peptide endothelial cell adhesion modulator #24.

KW Cadherin; cytostatic; gynecological; endometriosis;
 KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 KW cyclic.

XX Synthetic.

Key Location/Qualifiers

FT Modified-site 1 /note= "Linked to residue 6 to form a cyclic peptide,
 FT N-terminal acetyl"

FT Modified-site 6 /note= "Linked to residue 1 to form a cyclic peptide,
 FT C-terminal amide"

XX WO200177146-A2.

XX 18-OCT-2001.

XX 09-APR-2001; 2001WO-US11669.

XX 07-APR-2000; 2000US-0544782.

XX (UYMC-) UNIV MCGILL.

XX Blaschuk OW, Gour BJ, Farookhi R, Ali A;

XX WPI; 2002-049129/46.

XX Modulating endothelial cell adhesion for inhibiting development of
 PT endometriosis, increasing blood flow to tumor in a mammal, by
 PT contacting cell with a cyclic peptide having cadherin cell adhesion
 PT recognition sequence

Example 2; Page 50; 139pp; English.

XX The present invention relates to a method for modulating endothelial cell
 CC adhesion. The method comprises contacting an endothelial cell with a
 CC modulating peptide comprising a cadherin cell adhesion recognition
 CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 CC for inhibiting angiogenesis, increasing vasopermeability, increasing the
 CC blood flow to a tumour, disrupting neovasculture and inhibiting the
 CC development of endometriosis in a mammal. The modulating peptide reduces
 CC unwanted endothelial adhesion occurring between tumour cells, tumour
 CC cells and normal cells, normal cells as a result of surgery, injury,
 CC chemotherapy, disease and inflammation. The present sequence is one such
 CC modulating peptide.

SQ Sequence 6 AA;

Query Match 85.0%; Score 34; DB 23; Length 6;
 Best Local Similarity 83.3%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 |||||
 Db 1 CHAVSC 6

RESULT 17

ABG19097
 ID ABG19097 standard; Protein; 78 AA.

XX ABG19097;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19088.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS83284.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 20; SEQ ID No 49456; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase-chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 78 AA;

Query Match 85.0%; Score 34; DB 22; Length 78;
 Best Local Similarity 66.7%; Pred. No. 54;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 ||:|
 Db 11 CHSLDC 16

RESULT 18

ID ABB69902 standard; Protein; 649 AA.

XX AC ABB69902;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 36498.

XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL14005.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -

XX PS Disclosure; SEQ ID NO 36498; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 649 AA;

Query Match 85.0%; Score 34; DB 22; Length 649;
 Best Local Similarity 66.7%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 ||:|
 Db 26 CHIDC 31

RESULT 19

ID AAG65453 standard; peptide; 6 AA.

XX AC AAG65453;

XX DT 30-NOV-2001 (first entry)

XX DE Cyclic peptide with classical cadherin CAR sequence.

XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 XX cytostatic; vulnary; immunomodulator; vasotropic; neuroprotective;
 XX cerebroprotective; muscular; cyclic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 1 /note= "putative N-terminal acetylation or
 XX alkoxybenzylolation"

XX FT Modified-site 6 /note= "putative C-terminal amide or ester"

XX PN WO200153331-A2.

XX PD 26-JUL-2001.

XX PF 24-JAN-2001; 2001WO-US02508.

XX PR 24-JAN-2000; 2000US-0491078.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 XX Hu Z;

XX DR WPI; 2001-549899/61.

XX PT Cell adhesion modulating agent used for enhancing delivery of drug to
 XX tumor comprises imidazole compounds -

XX PS Examples; Page 415; 436pp; English.

XX CC The invention relates to cell adhesion modulating agents that comprise
 XX imidazole compounds of specified formulae that are peptidomimetics of
 XX cyclic peptides. The peptidomimetics have a structure similar to that of
 XX a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 XX The agents are used for modulating classical cadherin mediated
 XX intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 XX delivery of a drug to a tumor, inhibiting development of cancer,
 XX inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 XX wound healing, modulating the immune system, increasing vasopermeability,
 XX treating demyelinating disease, facilitating migration of an N-cadherin
 XX expressing cell on astrocytes, inhibiting synaptic stability, modulating
 XX neurite outgrowth, and treating spinal cord injuries and macular
 XX degeneration. The present sequence represents a cyclic peptide with
 XX classical cadherin cell adhesion recognition (CAR) sequence.

XX SQ Sequence 6 AA;

Query Match 82.5%; Score 33; DB 22; Length 6;
 Best Local Similarity 83.3%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6'
Db 1 CHAVPC 6

RESULT 20
AAM47548
ID AAM47548 standard; peptide; 6 AA.
XX AC AAM47548;
XX DT 12-FEB-2002 (first entry)
XX DE Cyclic peptide endothelial cell adhesion modulator #34.
XX KW Cadherin; cytostatic; gynecological; endometriosis;
XX KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
XX KW cyclic.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "Linked to residue 6 to form a cyclic peptide,
FT N-terminal acetyl"
FT Modified-site 6
FT /note= "Linked to residue 1 to form a cyclic peptide,
FT C-terminal amide"
XX WO200177146-A2.
XX PN 18-OCT-2001.
XX PD 09-APR-2001; 2001WO-US11669.
XX PF 07-APR-2000; 2000US-0544782.
XX PR (UYMC-) UNIV MCGILL.
XX PI Blaschuk OW, Gour BJ, Farookhi R, Ali A;
XX WPI; 2002-049129/06.
XX DR Modulating endothelial cell adhesion for inhibiting development of
XX PT endometriosis, increasing blood flow to tumor in a mammal, by
XX PT contacting cell with a cyclic peptide having cadherin cell adhesion
XX PT recognition sequence
XX PS Disclosure; Page 6; 139pp; English.
XX CC The present invention relates to a method for modulating endothelial cell
XX CC adhesion. The method comprises contacting an endothelial cell with a
XX CC modulating peptide comprising a cadherin cell adhesion recognition
XX CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
XX CC for inhibiting angiogenesis, increasing vasopermeability, increasing
XX CC blood flow to a tumour, disrupting neovasculation and inhibiting the
XX CC development of endometriosis in a mammal. The modulating peptide reduces
XX CC unwanted endothelial adhesion occurring between tumour cells, tumour
XX CC cells and normal cells, normal cells as a result of surgery, injury,
XX CC chemotherapy, disease and inflammation. The present sequence is one such
XX CC modulating peptide.
XX SQ Sequence 6 AA;

Query Match 82.5%; Score 33; DB 23; Length 6;
Best Local Similarity 83.3%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHAVPC 6

RESULT 21
AAM86916
ID AAM86916 standard; Protein; 85 AA.
XX AC AAM86916;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen SEQ ID NO:14509.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0190628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-MAR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226868.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234957.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 12-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI N-PSDB; AAKS9697.
DR WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis
XX
XX
PS Claim 11; SEQ ID NO 14509; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX
SQ Sequence 85 AA;
Query Match 82.5%; Score 33; DB 22; Length 85;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 CHAVDC 6
Db 4 CHSADC 9
||: ||
||: ||
RESULT 22
AAB12810
ID AAB12810 standard; Protein; 142 AA.
XX
XX AAB12810;
AC
XX
XX 27-NOV-2000 (first entry)
DE Mouse secretory type phospholipase A2 protein SEQ ID NO:15.
XX
XX Mouse; secretory phospholipase A2; PLA2; screening; development;
KW inhibitor.
XX
XX Mus musculus.
XX
XX JP2000166568-A.
XX
XX 20-JUN-2000.
PD

XX 09-DEC-1998; 98JP-0349604.
XX 09-DEC-1998; 98JP-0349604.
XX (SHIO) SHIONOGI & CO LTD.
XX WPI; 2000-485554/43.
XX N-PSDB; AAA73130.
XX New gene encoding mouse secretory type phospholipase A2 (PLA2) for
PT screening for inhibitors of PLA2 -
XX Claim 1; Page 13-14; 15pp; Japanese.
XX The present invention describes a mouse secretory type phospholipase A2
CC (PLA2) protein. The mouse secretory type PLA2-like protein can be used
CC for screening in the development of inhibitors against the function of
CC the protein. The present sequence represents mouse secretory type PLA2.
XX
SQ Sequence 142 AA;
Query Match 82.5%; Score 33; DB 21; Length 142;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CHAVDC 6
DB 62 CHAVDC 67
RESULT 23
ABG02071
ID ABG02071 standard; Protein; 142 AA.
XX
AC ABG02071;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #2062.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS66258.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 32430; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The

polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG00010-ABG30377 represent novel human
diagnostic amino acid sequences of the invention.
Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 142 AA;
Query Match 82.5%; Score 33; DB 22; Length 142;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CHAVDC 6
DB 25 CHAVDC 30
RESULT 24
AAB81021
ID AAB81021 standard; Protein; 142 AA.
XX
AC AAB81021;
XX
DT 12-JUN-2001 (first entry)
XX
DE Murine phospholipase A2 (PLA2) amino acid sequence.
XX
KW Phospholipase A2; PLA2; antibacterial; immunosuppressive; vulnery;
KW antinflammatory; tranquilliser; antiasthmatic; antiallergic; trauma;
KW antirheumatic; antiarthritic; septic shock; pancreatitis; mouse;
KW adult respiratory distress syndrome; ARDS; bronchial asthma;
KW allergic rhinitis; rheumatoid arthritis.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19 /label= Signal_peptide
FT Protein 20..142 /label= Mature_PLA2
FT /note= "Mature phospholipase A2"
XX
PN WO200121775-A1.
XX
XX 29-MAR-2001.
XX
PF 19-SEP-2000; 2000WO-JP06344.
XX
PR 21-SEP-1999; 99JP-0266616.
XX
PA (SHIO) SHIONOGI & CO LTD.
XX
PI Ishizaki J, Suzuki N, Hanasaki K;
XX WPI; 2001-290432/30.
DR N-PSDB; AAF77387.
XX
XX Human secretory phospholipase A2 and encoded gene, useful in diagnosis
PT of and screening drug candidates for treating associated diseases e.g.
PT septic shock, adult respiratory distress syndrome and rheumatoid
PT arthritis -

XX Example 3; Page 42; 50pp; Japanese.
PS This invention relates to human secretory phospholipase A2 (PLA2) protein
XX and the gene encoding it. Inhibitors of phospholipase A2 have
CC antibacterial; immunosuppressive; antiinflammatory; tranquilizer;
CC vulnary; antiasthmatic; antiallergic; antirheumatic; and antiarthritic
CC activity. The PLA2 protein, gene and an anti-PLA2 antibody are useful in
CC the diagnosis of PLA2 associated diseases e.g. septic shock, adult
CC allergic rhinitis and rheumatoid arthritis. The present sequence
CC represents human PLA2.
XX Sequence 142 AA;
SQ Query Match 82.5%; Score 33; DB 22; Length 142;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 CHAVDC 6
Db 62 CHAHC 67

RESULT 25
AAB81022
ID AAB81022 standard; Protein; 142 AA.
XX AAB81022;
XX 12-JUN-2001 (first entry)
DT Human phospholipase A2 (PLA2) amino acid sequence.
DE Phospholipase A2; PLA2; antibacterial; immunosuppressive; vulnary;
XX antiinflammatory; tranquilizer; antiasthmatic; antiallergic; trauma;
XX antirheumatic; antiarthritic; septic shock; pancreatitis; human;
XX adult respiratory distress syndrome; ARDS; bronchial asthma;
XX allergic rhinitis; rheumatoid arthritis.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..19
FT /label= Signal_peptide
FT Protein 20..142
FT /label= Mature_PLA2
FT /note= "Mature phospholipase A2"
XX WO200121775-A1.
XX 29-MAR-2001.
XX 18-SEP-2000; 2000WO-JP06344.
XX 21-SEP-1999; 99JP-0266616.
XX (SHIO) SHIONOGI & CO LTD..
XX Ishizaki J, Suzuki N, Hanasaki K;
XX WPI; 2001-290432/30.
XX N-PSDB; AAF77401.
XX Human secretory phospholipase A2 and encoded gene, useful in diagnosis
PT of and screening drug candidates for treating associated diseases e.g.
PT septic shock, adult respiratory distress syndrome and rheumatoid
PT arthritis
XX Claim 1; Page 46-47; 50pp; Japanese.
PS
XX

CC This invention relates to human secretory phospholipase A2 (PLA2) protein
CC and the gene encoding it. Inhibitors of phospholipase A2 have
CC antibacterial; immunosuppressive; antiinflammatory; tranquilizer;
CC vulnary; antiasthmatic; antiallergic; antirheumatic; and antiarthritic
CC activity. The PLA2 protein, gene and an anti-PLA2 antibody are useful in
CC the diagnosis of PLA2 associated diseases e.g. septic shock, adult
CC respiratory distress syndrome, pancreatitis, trauma, bronchial asthma,
CC allergic rhinitis and rheumatoid arthritis. The present sequence
CC represents human PLA2.
XX Sequence 142 AA;
SQ Query Match 82.5%; Score 33; DB 22; Length 142;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 CHAVDC 6
Db 62 CHAHC 67
Search completed: June 20, 2003, 20:52:01
Job time : 40.3182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 12.6818 Seconds
(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-20
Perfect score: 40
Sequence: 1 CHAVDC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	6	4	US-08-996-679-48
2	40	100.0	6	4	US-09-115-395-11
3	40	100.0	6	4	US-09-250-059-20
4	40	100.0	6	4	US-09-248-074-20
5	40	100.0	6	4	US-09-357-717-16
6	40	100.0	6	4	US-09-458-870-20
7	36	90.0	6	4	US-09-115-395-55
8	36	90.0	6	4	US-09-250-059-21
9	36	90.0	6	4	US-09-248-074-21
10	36	90.0	6	4	US-09-458-870-21
11	34	85.0	6	3	US-08-993-534A-14
12	34	85.0	6	4	US-08-996-679-14
13	34	85.0	6	4	US-09-115-395-42
14	34	85.0	6	4	US-09-507-102-14
15	34	85.0	6	4	US-09-250-059-38
16	34	85.0	6	4	US-09-248-074-38
17	34	85.0	6	4	US-09-357-717-25
18	34	85.0	6	4	US-09-458-870-38
19	33	82.5	6	4	US-09-458-870-99
20	32	80.0	226	1	US-07-828-798C-7
21	32	80.0	226	2	US-08-315-868A-7
22	32	80.0	226	3	US-08-495-819B-7
23	32	80.0	246	1	US-07-828-798C-6
24	32	80.0	246	2	US-08-315-868A-6
25	32	80.0	246	3	US-08-495-819B-6
26	31	77.5	6	4	US-09-458-870-94
27	31	77.5	7	3	US-08-893-534A-16

28	31	77.5	7	4	US-08-996-679-16	Sequence 16, Appl
29	31	77.5	7	4	US-09-115-395-13	Sequence 13, Appl
30	31	77.5	7	4	US-09-507-102-16	Sequence 16, Appl
31	31	77.5	7	4	US-09-250-059-26	Sequence 26, Appl
32	31	77.5	7	4	US-09-250-059-50	Sequence 50, Appl
33	31	77.5	7	4	US-09-248-074-26	Sequence 26, Appl
34	31	77.5	7	4	US-09-248-074-50	Sequence 50, Appl
35	31	77.5	7	4	US-09-357-717-19	Sequence 19, Appl
36	31	77.5	7	4	US-09-357-717-31	Sequence 31, Appl
37	31	77.5	7	4	US-09-458-870-26	Sequence 26, Appl
38	31	77.5	7	4	US-09-458-870-50	Sequence 50, Appl
39	31	77.5	7	4	US-08-996-679-50	Sequence 50, Appl
40	31	77.5	8	4	US-09-115-395-15	Sequence 15, Appl
41	31	77.5	8	4	US-09-250-059-28	Sequence 28, Appl
42	31	77.5	8	4	US-09-250-059-51	Sequence 51, Appl
43	31	77.5	8	4	US-09-248-074-28	Sequence 28, Appl
44	31	77.5	8	4	US-09-248-074-51	Sequence 51, Appl
45	31	77.5	8	4	US-09-357-717-20	Sequence 20, Appl
46	31	77.5	8	4	US-09-357-717-32	Sequence 32, Appl
47	31	77.5	8	4	US-09-458-870-28	Sequence 28, Appl
48	31	77.5	8	4	US-09-458-870-51	Sequence 51, Appl
49	31	77.5	9	4	US-08-996-679-52	Sequence 52, Appl
50	31	77.5	9	4	US-09-115-395-16	Sequence 16, Appl
51	31	77.5	9	4	US-09-250-059-32	Sequence 32, Appl
52	31	77.5	9	4	US-09-250-059-76	Sequence 76, Appl
53	31	77.5	9	4	US-09-248-074-32	Sequence 32, Appl
54	31	77.5	9	4	US-09-248-074-76	Sequence 76, Appl
55	31	77.5	9	4	US-09-357-717-22	Sequence 22, Appl
56	31	77.5	9	4	US-09-357-717-55	Sequence 55, Appl
57	31	77.5	9	4	US-09-458-870-32	Sequence 32, Appl
58	31	77.5	9	4	US-09-458-870-76	Sequence 76, Appl
59	31	77.5	846	1	US-08-356-354-2	Sequence 2, Appl
60	31	77.5	846	2	US-08-778-656-2	Sequence 2, Appl
61	31	77.5	908	1	US-08-356-354-6	Sequence 6, Appl
62	31	77.5	908	2	US-08-778-656-6	Sequence 6, Appl
63	31	77.5	1054	1	US-08-356-354-4	Sequence 4, Appl
64	31	77.5	1054	2	US-08-778-656-4	Sequence 4, Appl
65	31	77.5	3033	1	US-07-925-695-5	Sequence 5, Appl
66	30	75.0	6	3	US-08-893-534A-15	Sequence 15, Appl
67	30	75.0	6	4	US-08-996-679-15	Sequence 15, Appl
68	30	75.0	6	4	US-09-115-395-64	Sequence 64, Appl
69	30	75.0	6	4	US-09-507-102-15	Sequence 15, Appl
70	30	75.0	6	4	US-09-250-059-39	Sequence 39, Appl
71	30	75.0	6	4	US-09-248-074-39	Sequence 39, Appl
72	30	75.0	6	4	US-09-458-870-39	Sequence 39, Appl
73	30	75.0	487	1	US-08-218-943-2	Sequence 2, Appl
74	30	75.0	1140	2	US-08-657-641-7	Sequence 7, Appl
75	30	75.0	1140	5	PCT-US94-07233-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-996-679-48
; Sequence 48, Application US/08996679
; Patent No. 6169071
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gout, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-48

Query Match 100.0%; Score 40; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 2
US-09-115-395-11
; Sequence 11, Application US/09115395A
; Patent No. 6207639
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C3
; CURRENT APPLICATION NUMBER: US/09/115,395A
; EARLIER FILING DATE: 1998-07-14
; EARLIER FILING DATE: 1997-12-23
; EARLIER FILING DATE: 1997-07-11
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
; OTHER INFORMATION: Synthesis
US-09-115-395-11

Query Match 100.0%; Score 40; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 3
US-09-250-059-20
; Sequence 20, Application US/09250059
; Patent No. 633307
; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C6
; CURRENT APPLICATION NUMBER: US/09/250,059
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-250-059-20

Query Match 100.0%; Score 40; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 4
US-09-248-074-20
; Sequence 20, Application US/09248074
; Patent No. 6346512
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C5
; CURRENT APPLICATION NUMBER: US/09/248,074
; CURRENT FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-248-074-20

Query Match 100.0%; Score 40; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 5
US-09-357-717-16
; Sequence 16, Application US/09357717
; Patent No. 6417325

GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-357-717-16

Query Match 100.0%; Score 40; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 6
US-09-458-870-20
; Sequence 20, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Anmar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-20

Query Match 100.0%; Score 40; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 7
US-09-115-395-55
; Sequence 55, Application US/09115395A
; Patent No. 6207639
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C3
; CURRENT APPLICATION NUMBER: US/09/115,395A
; CURRENT FILING DATE: 1998-07-14
; EARLIER APPLICATION NUMBER: 08/996,679
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/893,534
; EARLIER FILING DATE: 1997-07-11
; EARLIER APPLICATION NUMBER: 60/021,612
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
; OTHER INFORMATION: Synthesis
US-09-115-395-55

Query Match 90.0%; Score 36; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHGVDC 6

RESULT 8
US-09-250-059-21
; Sequence 21, Application US/09250059
; Patent No. 633307
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C6
; CURRENT APPLICATION NUMBER: US/09/250,059
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: control peptide
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-250-059-21

Query Match 90.0%; Score 36; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHGVDC 6

RESULT 9

US-09-248-074-21
; Sequence 21, Application US/09248074
; Patent No. 6346512
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C5
; CURRENT APPLICATION NUMBER: US/09/248,074
; CURRENT FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: control peptide
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-248-074-21

Query Match 90.0%; Score 36; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHGVDC 6

RESULT 10

US-09-458-870-21
; Sequence 21, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: control peptide
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-21

Query Match 90.0%; Score 36; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHGVDC 6

RESULT 11

US-08-893-534A-14
; Sequence 14, Application US/08893534A
; Patent No. 6031072
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,534A
; FILING DATE: 11-JUL-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 100086.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6; amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
US-08-893-534A-14

Query Match 85.0%; Score 34; DB 3; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHAVSC 6

RESULT 12

US-08-996-679-14
; Sequence 14, Application US/08996679
; Patent No. 6169071
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-14

Query Match 85.0%; Score 34; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 1 CHAVSC 6

RESULT 13
US-09-115-395-42
Sequence 42, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115,395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996,679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893,534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis
US-09-115-395-42

Query Match 85.0%; Score 34; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 1 CHAVSC 6

RESULT 14
US-09-507-102-14
Sequence 14, Application US/09507102
Patent No. 6326352
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.

Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
CELL ADHESION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED IP LAW GROUP PLLC
STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,102
FILING DATE: 17-Feb-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/893,534
FILING DATE: 11-JUL-1997
APPLICATION NUMBER: US 60/021,612
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Christiansen, William T.
REGISTRATION NUMBER: 44,614
REFERENCE/DOCKET NUMBER: 100086.401C10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-507-102-14

Query Match 85.0%; Score 34; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 1 CHAVSC 6

RESULT 15
US-09-250-059-38
Sequence 38, Application US/09250059
Patent No. 6333307
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250,059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 38
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
FEATURE:

; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-250-059-38

Query Match 85.0%; Score 34; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
| | | | |
Db 1 CHAVSC 6

RESULT 16

US-09-248-074-38
; Sequence 38, Application US/09248074
; Patent No. 6346512
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C5
; CURRENT APPLICATION NUMBER: US/09/248,074
; CURRENT FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-248-074-38

Query Match 85.0%; Score 34; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
| | | | |
Db 1 CHAVSC 6

RESULT 17

US-09-357-717-25
; Sequence 25, Application US/09357717
; Patent No. 6417325
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion

; OTHER INFORMATION: recognition sequence

; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-357-717-25

Query Match 85.0%; Score 34; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
| | | | |
Db 1 CHAVSC 6

RESULT 18

US-09-458-870-38
; Sequence 38, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-38

Query Match 85.0%; Score 34; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
| | | | |
Db 1 CHAVSC 6

RESULT 19

US-09-458-870-99
; Sequence 99, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 6

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-458-870-99

Query Match 82.5%; Score 33; DB 4; Length 6;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
|||:
Db 1 CHAVPC 6

RESULT 20
US-07-828-798C-7
; Sequence 7, Application US/07828798C
; Patent No. 5389609
; GENERAL INFORMATION:
; APPLICANT: Woloshuk, Charles P.
; APPLICANT: Melchers, Leo S.
; APPLICANT: Cornelissen, Bernardus J. C.
; APPLICANT: Meulenhoff, Elisabeth J. S.
; APPLICANT: Sela-Buurlage, Marianne B.
; APPLICANT: Van Den Elzen, Petrus J. M.
; TITLE OF INVENTION: Antifungal Preparations, Process for
; TITLE OF INVENTION: Making Such Preparations, Process for Obtaining Plants with
; TITLE OF INVENTION: Decreased Susceptibility to Fungi
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61st Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM PC/XT/AT or compatibles
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/828,798C
; FILING DATE: 06-APR-1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL91/00089
; FILING DATE: 31-05-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30086
; REFERENCE/DOCKET NUMBER: U-8622-6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; TELEX: 233288
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 Amino Acids
; TYPE: Amino Acid
; TOPOLOGY: Linear

US-07-828-798C-7
Query Match 80.0%; Score 32; DB 1; Length 226;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
|||:
Db 141 CHAHC 146

RESULT 21
US-08-315-868A-7
; Sequence 7, Application US/08315868A
; Patent No. 5856151
; GENERAL INFORMATION:
; APPLICANT: Woloshuk, Charles P.
; APPLICANT: Melchers, Leo S.
; APPLICANT: Cornelissen, Bernardus J. C.
; APPLICANT: Meulenhoff, Elisabeth J. S.
; APPLICANT: Sela-Buurlage, Marianne B.
; APPLICANT: Van Den Elzen, Petrus J. M.
; TITLE OF INVENTION: Antifungal Preparations, Process for
; TITLE OF INVENTION: Making Such Preparations, Process for Obtaining Plants with
; TITLE OF INVENTION: Decreased Susceptibility to Fungi
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61st Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM PC/XT/AT or compatibles
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,868A
; FILING DATE: 09-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/828,798
; FILING DATE: 06-APR-1992
; APPLICATION NUMBER: PCT/NL91/00089
; FILING DATE: 31-05-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30086
; REFERENCE/DOCKET NUMBER: U-8622-6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; TELEX: 233288
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 Amino Acids
; TYPE: Amino Acid
; TOPOLOGY: Linear

US-08-315-868A-7
Query Match 80.0%; Score 32; DB 2; Length 226;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
|||:
Db 141 CHAHC 146

RESULT 22
US-08-495-819B-7
; Sequence 7, Application US/08495819B
; Patent No. 6087161
; GENERAL INFORMATION:
; APPLICANT: Woloshuk, Charles P.
; APPLICANT: Melchers, Leo S.

APPLICANT: Cornelissen, Bernardus J. C.
APPLICANT: Meulenhoff, Elisabeth J. S.
APPLICANT: Sela-Buurlage, Marianne B.
APPLICANT: Van Den Elzen, Petrus J. M.
TITLE OF INVENTION: Antifungal Preparations, Process for
TITLE OF INVENTION: Making Such Preparations, Process for Obtaining Plants
TITLE OF INVENTION: with Decreased Susceptibility to Fungi
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladas & Parry
STREET: 26 West 61st Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
COMPUTER: IBM PC/XT/AT or compatibles
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/495.819B
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/315,868
FILING DATE: 30-SEPT-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/828,798
FILING DATE: 06-APRIL-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL91/00089
FILING DATE: 31-05-91
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30086
REFERENCE/DOCKET NUMBER: U-010255-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
TELEFAX: (212) 246-8959
TELEX: 233288
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 Amino Acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-495-819B-7

Query Match 80.0%; Score 32; DB 3; Length 226;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
DB 141 CHAHC 146

RESULT 23
US-07-828-798C-6
Sequence 6, Application US/07828798C
Patent No. 5389609
GENERAL INFORMATION:
APPLICANT: Woloshuk, Charles P.
APPLICANT: Melchers, Leo S.
APPLICANT: Cornelissen, Bernardus J. C.
APPLICANT: Meulenhoff, Elisabeth J. S.
APPLICANT: Sela-Buurlage, Marianne B.
APPLICANT: Van Den Elzen, Petrus J. M.
TITLE OF INVENTION: Antifungal Preparations, Process for
TITLE OF INVENTION: Making Such Preparations, Process for Obtaining Plants with
TITLE OF INVENTION: Decreased Susceptibility to Fungi
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ladas & Parry
STREET: 26 West 61st Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
COMPUTER: IBM PC/XT/AT or compatibles
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/828,798C
FILING DATE: 06-APR-1992
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL91/00089
FILING DATE: 31-05-91
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30086
REFERENCE/DOCKET NUMBER: U-8622-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
TELEFAX: (212) 246-8959
TELEX: 233288
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 246 Amino Acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-07-828-798C-6

Query Match 80.0%; Score 32; DB 1; Length 246;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
DB 141 CHAHC 146

RESULT 24
US-08-315-868A-6
Sequence 6, Application US/08315868A
Patent No. 5856151
GENERAL INFORMATION:
APPLICANT: Woloshuk, Charles P.
APPLICANT: Melchers, Leo S.
APPLICANT: Cornelissen, Bernardus J. C.
APPLICANT: Meulenhoff, Elisabeth J. S.
APPLICANT: Sela-Buurlage, Marianne B.
APPLICANT: Van Den Elzen, Petrus J. M.
TITLE OF INVENTION: Antifungal Preparations, Process for
TITLE OF INVENTION: Making Such Preparations, Process for Obtaining Plants with
TITLE OF INVENTION: Decreased Susceptibility to Fungi
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladas & Parry
STREET: 26 West 61st Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
COMPUTER: IBM PC/XT/AT or compatibles
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,868A
FILING DATE: 09-SEP-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/828,798
; FILING DATE: 06-APR-1992
; APPLICATION NUMBER: PCT/NL91/00089
; FILING DATE: 31-03-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30086
; REFERENCE/DOCKET NUMBER: U-8622-6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; TELEX: 233288
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 Amino Acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-315-868A-6

Query Match 80.0%; Score 32; DB 2; Length 246;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 141 CHAHC 146

RESULT 25

US-08-495-819B-6
; Sequence 6, Application US/08495819B
; Patent No. 6087161

GENERAL INFORMATION:

; APPLICANT: Woloshuk, Charles P.
; APPLICANT: Melchers, Leo S.
; APPLICANT: Cornelissen, Bernardus J. C.
; APPLICANT: Meulenhoff, Elisabeth J. S.
; APPLICANT: Sela-Buurlage, Marianne B.
; APPLICANT: Van Den Elzen, Petrus J. M.
; TITLE OF INVENTION: Antifungal Preparations, Process for
; TITLE OF INVENTION: Making Such Preparations, Process for Obtaining Plants
; TITLE OF INVENTION: With Decreased Susceptibility to Fungi
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61st Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM PC/XT/AT or compatibles
; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/495,819B

; FILING DATE: 31-MAY-1995

; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/315,868

; FILING DATE: 30-SEPT-94

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/828,798

; FILING DATE: 06-APRIL-92

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/NL91/00089

; FILING DATE: 31-05-91

ATTORNEY/AGENT INFORMATION:

; NAME: Mass, Clifford J.

; REGISTRATION NUMBER: 30086

; REFERENCE/DOCKET NUMBER: U-010255-1

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; TELEX: 233288
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 Amino Acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-495-819B-6

Query Match 80.0%; Score 32; DB 3; Length 246;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 141 CHAHC 146

Search completed: June 20, 2003, 21:00:33
Job time : 12.6818 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:52:19 ; Search time 15.4091 Seconds
(without alignments)
42.134 Million cell updates/sec

Title: US-10-105-008-20
Perfect score: 40
Sequence: 1 CHAVDC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pdb.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pdb.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pdb.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pdb.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pdb.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pdb.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pdb.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	6	9	US-09-769-145-20
2	40	100.0	6	9	US-10-058-821-16
3	40	100.0	6	10	US-09-234-395-297
4	40	100.0	6	10	US-09-305-928-297
5	36	90.0	6	9	US-09-769-145-21
6	35	87.5	1798	10	US-09-938-275-9
7	35	87.5	1798	10	US-09-845-583-8
8	34	85.0	6	9	US-09-769-145-38
9	34	85.0	6	9	US-10-058-821-25
10	34	85.0	6	10	US-09-234-395-305
11	34	85.0	6	10	US-09-305-928-305
12	34	85.0	6	12	US-10-006-982-14
13	33	82.5	6	9	US-09-769-145-93
	33	82.5	154	9	US-09-975-456B-8
	33	82.5	168	9	US-09-975-456B-2
	33	82.5	133	10	US-09-764-864-1497
	33	82.5	1248	9	US-10-108-605-59
			6	9	US-09-769-145-91
			7	9	US-09-769-145-26

Mon

20	31	77.5	7	9	US-09-769-145-50	Sequence 50, Appl
21	31	77.5	7	9	US-10-058-821-19	Sequence 19, Appl
22	31	77.5	7	9	US-10-058-821-31	Sequence 31, Appl
23	31	77.5	7	10	US-09-234-395-299	Sequence 299, App
24	31	77.5	7	10	US-09-305-928-299	Sequence 299, App
25	31	77.5	7	12	US-10-006-982-16	Sequence 16, Appl
26	31	77.5	8	9	US-09-769-145-28	Sequence 28, Appl
27	31	77.5	8	9	US-09-769-145-51	Sequence 51, Appl
28	31	77.5	8	9	US-10-058-821-20	Sequence 20, Appl
29	31	77.5	8	9	US-10-058-821-32	Sequence 32, Appl
30	31	77.5	8	10	US-09-234-395-301	Sequence 301, App
31	31	77.5	8	10	US-09-305-928-301	Sequence 301, App
32	31	77.5	9	9	US-09-769-145-32	Sequence 32, Appl
33	31	77.5	9	9	US-09-769-145-76	Sequence 76, Appl
34	31	77.5	9	9	US-10-058-821-22	Sequence 22, Appl
35	31	77.5	9	9	US-10-058-821-55	Sequence 55, Appl
36	31	77.5	9	10	US-09-234-395-303	Sequence 303, App
37	31	77.5	9	10	US-09-305-928-303	Sequence 303, App
38	31	77.5	81	10	US-09-735-712-4	Sequence 4, Appl
39	31	77.5	149	10	US-09-735-712-8	Sequence 8, Appl
40	31	77.5	199	10	US-09-735-712-2	Sequence 2, Appl
41	31	77.5	200	9	US-09-924-340-34	Sequence 34, Appl
42	31	77.5	200	9	US-09-924-600A-34	Sequence 34, Appl
43	31	77.5	200	9	US-10-000-489-34	Sequence 34, Appl
44	31	77.5	200	9	US-10-000-986-34	Sequence 34, Appl
45	31	77.5	200	10	US-09-821-821-2	Sequence 2, Appl
46	31	77.5	200	10	US-09-731-872-258	Sequence 258, App
47	31	77.5	646	9	US-10-062-937B-10	Sequence 10, Appl
48	31	77.5	846	10	US-09-376-045-2	Sequence 2, Appl
49	31	77.5	908	10	US-09-376-045-6	Sequence 6, Appl
50	31	77.5	1053	9	US-10-217-700-6	Sequence 6, Appl
51	31	77.5	1054	10	US-09-376-045-4	Sequence 4, Appl
52	30	75.0	6	9	US-09-769-145-39	Sequence 39, Appl
53	30	75.0	6	12	US-10-006-982-15	Sequence 15, Appl
54	30	75.0	183	10	US-09-764-864-1517	Sequence 1517, Ap
55	30	75.0	197	10	US-09-764-864-1097	Sequence 1097, Ap
56	30	75.0	280	10	US-09-823-356-7	Sequence 7, Appl
57	30	75.0	475	10	US-09-815-242-11259	Sequence 11259, A
58	30	75.0	1056	9	US-10-217-700-1	Sequence 1, Appl
59	30	75.0	1172	9	US-09-974-298-56	Sequence 56, Appl
60	30	75.0	1172	9	US-10-299-058-10	Sequence 10, Appl
61	30	75.0	1172	10	US-09-919-172-16	Sequence 16, Appl
62	30	75.0	1940	9	US-10-016-283-34	Sequence 34, Appl
63	29	72.5	12	9	US-10-158-847-99	Sequence 99, Appl
64	29	72.5	17	9	US-09-910-009A-226	Sequence 226, App
65	29	72.5	57	9	US-09-948-783-207	Sequence 207, App
66	29	72.5	58	9	US-09-892-877-206	Sequence 206, App
67	29	72.5	155	9	US-09-975-456B-10	Sequence 10, Appl
68	29	72.5	165	9	US-10-124-591-1	Sequence 1, Appl
69	29	72.5	227	9	US-10-108-605-17	Sequence 17, Appl
70	29	72.5	406	9	US-09-839-446-20	Sequence 20, Appl
71	29	72.5	406	10	US-09-898-570-20	Sequence 20, Appl
72	29	72.5	411	10	US-09-854-731-22	Sequence 22, Appl
73	29	72.5	524	9	US-10-108-605-359	Sequence 359, App
74	29	72.5	1059	10	US-09-800-729-217	Sequence 217, App
75	29	72.5	2050	9	US-09-839-446-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1:
US-09-769-145-20
; Sequence 20, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Anmar
; APPLICANT: Ali, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoameng

APPLICANT: Hu, Zenglian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR FILING DATE: 2000-01-24
PRIORITY FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-20

Query Match 100.0%; Score 40; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 2
US-10-058-821-16
Sequence 16, Application US/10058821
Publication No. US2003008781A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
APPLICANT: Farookhi, Riaz
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
FILE REFERENCE: 100086.401C12
CURRENT APPLICATION NUMBER: US/10/058,821
CURRENT FILING DATE: 2002-01-29
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-10-058-821-16

Query Match 100.0%; Score 40; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 3
US-09-234-395-297
Sequence 297, Application US/09234395
Patent No. US20020123044A1

GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 297
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-234-395-297

Query Match 100.0%; Score 40; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 4
US-09-305-928-297
Sequence 297, Application US/09305928
Patent No. US20020146687A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C4
CURRENT APPLICATION NUMBER: US/09/305,928
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 297
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: N-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-305-928-297

Query Match 100.0%; Score 40; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 1 CHAVDC 6

RESULT 5
US-09-769-145-21
Sequence 21, Application US/09769145
Patent No. US20020168761A1
GENERAL INFORMATION:
APPLICANT: Gour, Barbara J.
APPLICANT: Blaschuk, Orest W.

APPLICANT: Ali, Annar
APPLICANT: Ni, Feng
APPLICANT: Chen, Zhigang
APPLICANT: Michaud, Stephanie
APPLICANT: Wang, Shaoameng
APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 21
LENGTH: 6
TYPE: PRT
ORGANISM: Homo sapiens
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: control peptide
OTHER INFORMATION: Cyclic peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-21

Query Match 90.0%; Score 36; DB 9; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHGVDC 6

RESULT 6
US-09-938-275-9
Sequence 9, Application US/09938275
Patent No. US20020111309A1
GENERAL INFORMATION:
APPLICANT: Gerardo Castillo
APPLICANT: Alan Snow
TITLE OF INVENTION: Therapeutic and Diagnostic Applications
TITLE OF INVENTION: of Laminin and Laminin-Derived Protein Fragments
FILE REFERENCE: PROTEO.P03
CURRENT APPLICATION NUMBER: US/09/938,275
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 1798
TYPE: PRT
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Swissprot P55268
DATABASE ENTRY DATE: 1996-10-01
US-09-938-275-9

Query Match 87.5%; Score 35; DB 10; Length 1798;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
1140 CHACDC 1145

Mon Jun 23 10:09:31 2003

US/09845583

APPLICANT: Burgeson, Robert
APPLICANT: Brunken, William Joseph
APPLICANT: Champilaud, Marie-France
APPLICANT: Hunter, Dale
TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
FILE REFERENCE: 10287-056001
CURRENT APPLICATION NUMBER: US/09/845,583
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: US 60/200,863
PRIOR FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 1798
TYPE: PRT
ORGANISM: Homo sapiens
ORGANISM: Artificial Sequence
US-09-845-583-8

Query Match 87.5%; Score 35; DB 10; Length 1798;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1140 CHACDC 1145

RESULT 8
US-09-769-145-38
Sequence 38, Application US/09769145
Patent No. US20020168761A1
GENERAL INFORMATION:
APPLICANT: Gour, Barbara J.
APPLICANT: Blaschuk, Orest W.
APPLICANT: Ali, Annar
APPLICANT: Ni, Feng
APPLICANT: Chen, Zhigang
APPLICANT: Michaud, Stephanie
APPLICANT: Wang, Shaoameng
APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 38
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-38

Query Match 85.0%; Score 34; DB 9; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHAVSC 6

RESULT 9
US-10-058-821-25

; Sequence 25, Application US/10058821
; Publication No. US20030087811A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farcokhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C12
; CURRENT APPLICATION NUMBER: US/10/058,821
; CURRENT FILING DATE: 2002-01-29
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-10-058-821-25

Query Match 85.0%; Score 34; DB 9; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHAVSC 6

RESULT 10
US-09-234-395-305
; Sequence 305, Application US/09234395
; Patent No. US20020123044A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
; FILE REFERENCE: 100086.407C2
; CURRENT APPLICATION NUMBER: US/09/234,395
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 305
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: Synthesis and Cyclization based on Human
; OTHER INFORMATION: N-Cadherin
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-234-395-305

Query Match 85.0%; Score 34; DB 10; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHAVSC 6

RESULT 11
US-09-305-928-305
; Sequence 305, Application US/09305928

; Patent No. US20020146687A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
; FILE REFERENCE: 100086.407C4
; CURRENT APPLICATION NUMBER: US/09/305,928
; CURRENT FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 305
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: Synthesis and Cyclization based on Human
; OTHER INFORMATION: N-Cadherin
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-305-928-305

Query Match 85.0%; Score 34; DB 10; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 1 CHAVSC 6

RESULT 12
US-10-006-982-14
; Sequence 14, Application US/10006982
; Patent No. US20020151475A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; CELL ADHESION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED IP LAW GROUP PLLC
; STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/006,982
; FILING DATE: 04-Dec-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 100086.401C11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-10-006-982-14

Query Match 85.0%; Score 34; DB 12; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
| | | | |
DB 1 CHAVSC 6

RESULT 13

US-09-769-145-93
; Sequence 93, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Annmar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shoameng
; APPLICANT: Hu, Zengjian
; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
; FILE REFERENCE: 100086.413C1
; CURRENT APPLICATION NUMBER: US/09/769,145
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 09/491,078
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 93
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
; OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-769-145-93

Query Match 82.5%; Score 33; DB 9; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
| | | | |
DB 1 CHAVPC 6

RESULT 14

US-09-975-456B-8
; Sequence 8, Application US/09975456B
; Publication No. US20030073087A1
; GENERAL INFORMATION:
; APPLICANT: LAZDUNSKI, MICHEL
; APPLICANT: LAMBEAU, GERARD
; APPLICANT: VALENTIN, EMMANUEL
; TITLE OF INVENTION: NOVEL MAMMALIAN SECRETED GROUP IIF PHOSPHOLIPASE A2
; FILE REFERENCE: 1478-R-00
; CURRENT APPLICATION NUMBER: US/09/975,456B
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 60/239,491
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NOS: 10
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (305)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (311)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (315)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Mon Jun 23 10:09:31 2003

; ORGANISM: Homo sapiens
US-09-975-456B-8

Query Match 82.5%; Score 33; DB 9; Length 154;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
| | | | |
DB 74 CHAVDC 79

RESULT 15

US-09-975-456B-2
; Sequence 2, Application US/09975456B
; Publication No. US20030073087A1
; GENERAL INFORMATION:
; APPLICANT: LAZDUNSKI, MICHEL
; APPLICANT: LAMBEAU, GERARD
; APPLICANT: VALENTIN, EMMANUEL
; TITLE OF INVENTION: NOVEL MAMMALIAN SECRETED GROUP IIF PHOSPHOLIPASE A2
; FILE REFERENCE: 1478-R-00
; CURRENT APPLICATION NUMBER: US/09/975,456B
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 60/239,491
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 2.1
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-975-456B-2

Query Match 82.5%; Score 33; DB 9; Length 168;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
| | | | |
DB 64 CHAVDC 69

RESULT 16

US-09-764-864-1497
; Sequence 1497, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1497
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (305)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (311)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (315)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE
LOCATION: (316)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (317)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (322)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (329)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1497

Query Match 82.5%; Score 33; DB 9; Length 1248;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 39 CHAGDC 44

RESULT 17
US-10-108-605-59
Sequence 59, Application US/10108605
Patent No. US20020160934A1
GENERAL INFORMATION:
APPLICANT: Broadus, Julie
APPLICANT: Stam, Lynn
APPLICANT: Bachmann, Jane
APPLICANT: Kamdar, Kim
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
FILE REFERENCE: 31133B
CURRENT APPLICATION NUMBER: US/10/108,605
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: US 09/761,142
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/176,418
PRIOR FILING DATE: 2000-01-14
NUMBER OF SEQ ID NOS: 361
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 59
LENGTH: 1248
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-10-108-605-59

Query Match 82.5%; Score 33; DB 9; Length 1248;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 673 CHSADC 678

RESULT 18
US-09-769-145-91
Sequence 91, Application US/09769145
Patent No. US20020168761A1
GENERAL INFORMATION:
APPLICANT: Gour, Barbara J.
APPLICANT: Blaschuk, Orest W.
APPLICANT: Ali, Annmar
APPLICANT: Ni, Feng
APPLICANT: Chen, Zhigang
APPLICANT: Michaud, Stephanie
APPLICANT: Wang, Shoameng
APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1

CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 91
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic Peptide
OTHER INFORMATION: with Classical Cell Adhesion Recognition Sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-91

Query Match 77.5%; Score 31; DB 9; Length 6;
Best Local Similarity 83.3%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 1 CHAVYC 6

RESULT 19
US-09-769-145-26
Sequence 26, Application US/09769145
Patent No. US20020168761A1
GENERAL INFORMATION:
APPLICANT: Gour, Barbara J.
APPLICANT: Blaschuk, Orest W.
APPLICANT: Ali, Annmar
APPLICANT: Ni, Feng
APPLICANT: Chen, Zhigang
APPLICANT: Michaud, Stephanie
APPLICANT: Wang, Shoameng
APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-769-145-26

Query Match 77.5%; Score 31; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVDC 6
Db 3 HAVDC 7

RESULT 20

US-09-769-145-50
; Sequence 50, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Anmar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shaoenag
; APPLICANT: Hu, Zengjian
; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
; FILE REFERENCE: 100086.413C1
; CURRENT APPLICATION NUMBER: US/09/769,145
; CURRENT FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 09/491,078
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-769-145-50

Query Match 77.5%; Score 31; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVD 5
DB 1 CHAVD 5

RESULT 21
US-10-058-821-19
; Sequence 19, Application US/10058821
; Publication No. US2003008781A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C12
; CURRENT APPLICATION NUMBER: US/10/058,821
; CURRENT FILING DATE: 2002-01-29
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-10-058-821-19

Mon Jun 23 10:09:31 2003

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HAVDC 6
DB 3 HAVDC 7

RESULT 22
US-10-058-821-31
; Sequence 31, Application US/10058821
; Publication No. US2003008781A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C12
; CURRENT APPLICATION NUMBER: US/10/058,821
; CURRENT FILING DATE: 2002-01-29
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-10-058-821-31

Query Match 77.5%; Score 31; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVD 5
DB 1 CHAVD 5

RESULT 23
US-09-234-395-299
; Sequence 299, Application US/09234395
; Patent No. US20020123044A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
; FILE REFERENCE: 100086.407C2
; CURRENT APPLICATION NUMBER: US/09/234,395
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 299
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: Synthesis and Cyclization based on Human
; OTHER INFORMATION: N-Cadherin
; FEATURE:
; OTHER INFORMATION: Cyclic-P
US-09-234-395-299

QY 2 HAVDC 6
Db 3 HAVDC 7

RESULT 24

US-09-305-928-299
; Sequence 299, Application US/09305928
; Patent No. US20020146687A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
; FILE REFERENCE: 100086.407C4
; CURRENT APPLICATION NUMBER: US/09/305,928
; CURRENT FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 299
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: Synthesis and Cyclization based on Human
; OTHER INFORMATION: N-Cadherin
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-305-928-299

Query Match 77.5%; Score 31; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVDC 6
Db 3 HAVDC 7

RESULT 25

US-10-006-982-16
; Sequence 16, Application US/10006982
; Patent No. US20020151475A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; CELL ADHESION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED IP LAW GROUP PLLC
; STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/006,982
; FILING DATE: 04-Dec-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 100086.401C11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-006-982-16

Query Match 77.5%; Score 31; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HAVDC 6
Db 3 HAVDC 7

Search completed: June 20, 2003, 21:02:43
Job time : 21.4091 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 15.2727 Seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-20
Perfect score: 40
Sequence: 1 CHAVDC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

PIR 73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	37	92.5	225	1	WZBEF1		gene 58 protein -
2	36	90.0	169	2	A37390		hypothetical 19K p
3	36	90.0	169	2	S07015		X polypeptide (ORF
4	36	90.0	206	2	T24484		hypothetical prote
5	36	90.0	589	2	T43210		fibulin-ID precurs
6	36	90.0	689	2	T42760		fibulin, splice fo
7	36	90.0	712	2	T42990		fibulin 1, splice
8	36	90.0	798	2	T22793		hypothetical prote
9	36	90.0	1275	2	T41523		hypothetical rhol
10	36	90.0	1404	2	T19277		hypothetical prote
11	35	87.5	1168	2	I56985		kalinin B1 - mouse
12	35	87.5	1797	2	A55677		laminin beta-2 cha
13	35	87.5	1798	2	S33869		laminin beta-2 cha
14	34	85.0	268	2	A25394		enoyl-(acyl-carrie
15	34	85.0	270	2	A57376		enoyl-(acyl-carrie
16	34	85.0	354	2	A56609		uroporphyrinogen d
17	33	82.5	130	2	D83305		hypothetical prote
18	33	82.5	1188	2	D86236		protein Fl4N23.5 [
19	32	80.0	152	2	D55961		hypothetical prote
20	32	80.0	233	2	S31829		osmotin-like relat
21	32	80.0	238	2	S28001		osmotin-like relat
22	32	80.0	239	2	S07406		thauromatin homolog
23	32	80.0	246	2	S30144		osmotin-like prote
24	32	80.0	247	2	S33197		osmotin-like prote
25	32	80.0	247	2	S33156		osmotin-like prote
26	32	80.0	250	2	S30157		osmotin precursor
27	32	80.0	302	2	G86260		protein T12C24.28
28	32	80.0	420	2	S43559		coiled coil protei
29	32	80.0	429	2	T21113		hypothetical prote

30	32	80.0	454	2	T05567		allin lyase homol
31	32	80.0	612	2	B42755		E-selectin precurs
32	32	80.0	660	1	UYFVFP		noncapsid protein
33	32	80.0	662	1	UYPVNA		noncapsid protein
34	31	77.5	38	2	H81603		hypothetical prote
35	31	77.5	127	2	A49269		phospholipase A2 h
36	31	77.5	200	2	JC7585		testis expressed t
37	31	77.5	219	2	S18843		TFG1 protein - yea
38	31	77.5	233	2	S52407		TATA box-binding p
39	31	77.5	254	2	AH0249		probable copper ho
40	31	77.5	270	2	G90272		conserved hypoteth
41	31	77.5	286	2	E65103		tagatose-bisphosph
42	31	77.5	286	2	A91131		tagatose-1,6-bisph
43	31	77.5	286	2	A85976		tagatose-bisphosph
44	31	77.5	299	2	S75513		protoporphyrinogen
45	31	77.5	304	2	AC1821		protoporphyrinogen
46	31	77.5	342	2	T35500		6-phosphofructokin
47	31	77.5	392	1	UYPVIF		noncapsid protein
48	31	77.5	396	2	F84326		GTP-binding protei
49	31	77.5	416	2	H69252		alcohol dehydrogen
50	31	77.5	431	2	G81179		oxidoreductase, pr
51	31	77.5	508	2	F86458		unknown protein, 7
52	31	77.5	516	2	H82973		choline transporte
53	31	77.5	601	2	T35962		probable oxidoredu
54	31	77.5	626	2	D88601		protein Y49B10.11
55	31	77.5	668	1	UYPVCP		noncapsid protein
56	31	77.5	668	1	UYPVME		noncapsid protein
57	31	77.5	668	1	UYPVFP		noncapsid protein
58	31	77.5	799	2	T48690		hypothetical prote
59	31	77.5	1053	2	S34172		sucrose-phosphate
60	31	77.5	1378	2	T47605		RING finger-like p
61	31	77.5	1424	2	S11480		hypothetical prote
62	31	77.5	1441	2	T00335		hypothetical prote
63	31	77.5	2048	1	ZLNZSE		genome polypotein
64	31	77.5	2228	1	ZLNZSV		genome polypotein
65	31	77.5	3033	1	JQ1303		cytochrome c552 -
66	30	75.0	60	2	JN0251		cytochrome c551 -
67	30	75.0	82	1	CCAV95		hypothetical prote
68	30	75.0	95	2	G97394		conserved hypoteth
69	30	75.0	95	2	AI2612		cytochrome c551 pr
70	30	75.0	104	1	CCP85A		hypothetical 12.4K
71	30	75.0	124	2	J50515		hypothetical prote
72	30	75.0	155	2	T26776		multiple resistanc
73	30	75.0	159	2	D97470		Na+/H+ antiporter
74	30	75.0	159	2	AI2688		gene 58 protein -
75	30	75.0	227	2	T42601		

ALIGNMENTS

RESULT 1

WZBEF1

gene 58 protein - equine herpesvirus 1 (strain Ab4p)

C:Species: equine herpesvirus 1

A:Note: host Equus caballus (domestic horse)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: D36801

R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to GenBank, March 1992

A:Description: The DNA sequence of equine herpesvirus-1.

A:Reference number: A36805

A:Accession: D36801

A:Molecule type: DNA

A:Residues: 1-225 <TEL>

A:Cross-references: GB:M86664; NID:G330791; PIDN:AB02493.1; PID:G330849

R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

Virology 189, 304-316, 1992

A:Title: The DNA sequence of equine herpesvirus-1.

A:Reference number: A41831; MUID:92295566; PMID:1318606

A:Contents: annotation; possible protein-coding frames

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 58
C:Superfamily: varicella-zoster virus gene 56 protein

Query Match 92.5%; Score 37; DB 1; Length 225;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 149 CHALDC 154

RESULT 2

A37390
hypothetical 19K protein (orf 5' region) - Escherichia coli plasmid P307
C:Species: Escherichia coli
C>Date: 17-Apr-1993 #sequence_revision 14-May-1993 #text_change 10-Dec-1999
C:Accession: A37390
R:Graus-Goeldner, A.; Graus, H.; Schlacher, T.; Hoegenauer, G.
Plasmid 24, 119-131, 1990
A:Title: The sequences of genes bordering orf in the enterotoxin plasmid P307: comparison
A:Reference number: A37390; MUID:91261994; PMID:2096398
A:Accession: A37390
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-169 <GRA>
A:Cross-references: GB:M62986; NID:gl50463; PIDN:AA25521.1; PID:gl50464
A:Note: the authors translated the codon ACG for residue 58 as Ile
C:Genetics:
A:Genome: plasmid
C:Superfamily: Escherichia coli gene X protein

Query Match 90.0%; Score 36; DB 2; Length 169;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 17 CHATDC 22

RESULT 3

S07015
X polypeptide (ORF 19) (P19 protein) (ORF169) - Escherichia coli plasmid F
C:Species: Escherichia coli
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 27-Nov-2001
C:Accession: S07015
R:Loth, S.; Cram, D.; Skurray, R.
Mol. Gen. Genet. 219, 177-186, 1989
A:Title: Nucleotide sequence of the leading region adjacent to the origin of transfer on
A:Reference number: S07013; MUID:90136505; PMID:2693941
A:Accession: S07015
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-169 <LOH>
A:Cross-references: GB:X17539; NID:g41912; PIDN:CRA35576.1; PID:g41914
C:Genetics:
A:Genome: plasmid
C:Superfamily: Escherichia coli gene X protein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-169/Product: hypothetical protein 169 #status predicted <MAT>

Query Match 90.0%; Score 36; DB 2; Length 169;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 17 CHATDC 22

RESULT 4

T24484

hypothetical protein T05A1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24484
R:Lloyd, C.
Submitted to the EMBL Data Library, December 1995
A:Reference number: Z19897
A:Accession: T24484
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-206 <WIL>
A:Cross-references: EMBL:Z68219; PIDN:CAA92478.1; GSPDB:GN00022; CESP:T05A1.6
A:Experimental source: clone T05A1
C:Genetics:
A:Gene: CESP:T05A1.6
A:Map position: 4
A:Introns: 57/2; 126/1

Query Match 90.0%; Score 36; DB 2; Length 206;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 40 CHSIDC 45

RESULT 5

T43210
fibulin-1D precursor - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T43210
R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.
Submitted to the EMBL Data Library, June 1998
A:Description: Identification of chicken and C. elegans fibulin-1 homologs and chara
A:Reference number: Z22337
A:Accession: T43210
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-589 <BAR>
A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1
C:Genetics:
A:Note: Intron positions not resolved (incomplete sequence)
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 90.0%; Score 36; DB 2; Length 589;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 446 CHSIDC 451

RESULT 6

T42760
fibulin, splice form D precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42760
R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.
Submitted to the EMBL Data Library, February 1998
A:Description: Isolation of chicken and nematode fibulin-1 homologs and characteriza
A:Reference number: Z22267
A:Accession: T42760
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-689 <BAR>
A:Cross-references: EMBL:AF051401; PIDN:AAC28321.1
C:Genetics:
A:Note: PBLN1
C:Superfamily: unassigned EGF-related proteins; EGF homology

0;

Query Match 90.0%; Score 36; DB 2; Length 689;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
||:|
Db - 546 CHSIDC 551

RESULT 7

T42990
C:Species: Caenorhabditis elegans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42990
R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.
submitted to the EMBL Data Library, February 1998
A:Description: Isolation of chicken and nematode fibulin-1 homologs and characterization
A:Reference number: Z22267
A:Accession: T42990
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-712 <BAR>
A:Cross-references: EMBL:AF051402; PIDN:AAC28322.1
C:Genetics:
A:Gene: FBLN1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; basement membrane; extracellular matrix

Query Match 90.0%; Score 36; DB 2; Length 712;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
||:|
Db - 546 CHSIDC 551

RESULT 8

T22793
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Aug-2002
C:Accession: T22793; T24489
R:Wilkinson, J.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19616
A:Accession: T22793
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-798 <WIL>
A:Cross-references: EMBL:Z68749; PIDN:CAA92962.1; GSPDB:GN000022; CESP:F56H11.1
A:Experimental source: clone F56H11
R:Lloyd, C.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z19897
A:Accession: T24489
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-798 <WIL>
A:Cross-references: EMBL:Z68219; PIDN:CAA92483.1; GSPDB:GN000022; CESP:F56H11.1
A:Experimental source: clone T05A1
C:Genetics:
A:Gene: CESP:F56H11.1
A:Map position: 4
A:Introns: 14/1; 92/3; 144/1; 195/1; 281/1; 325/2; 371/2; 390/1; 437/1; 471/3; 498/3; 60
C:Superfamily: fibulin-1; EGF homology

Query Match 90.0%; Score 36; DB 2; Length 798;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
||:|
Db - 632 CHSIDC 637

RESULT 9

T41523
hypothetical rhol gdp-gtp exchange protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T41523
R:Wood, V.; Rajadream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z22000
A:Accession: T41523
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1275 <WOO>
A:Cross-references: EMBL:AL049498; PIDN:CAB39902.1; GSPDB:GN000068; SPDB:SPCC645.06c
A:Experimental source: strain 972h-; cosmid c645
C:Genetics:
A:Gene: SPDB:SPCC645.06c
A:Map position: 3

Query Match 90.0%; Score 36; DB 2; Length 1275;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
||:|
Db - 1099 CHSIDC 1104

RESULT 10

T19277
hypothetical protein C14B4.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T19277; T19909
R:Basham, V.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19100
A:Accession: T19277
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1404 <WIL>
A:Cross-references: EMBL:Z81471; PIDN:CAB03887.1; GSPDB:GN000023; CESP:C14B4.2
A:Experimental source: clone C14B4
R:Basham, V.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19194
A:Accession: T19909
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1404 <WIL>
A:Cross-references: EMBL:Z81483; PIDN:CAB03966.1; GSPDB:GN000023; CESP:C14B4.2
A:Experimental source: clone C43D7
C:Genetics:
A:Gene: CESP:C14B4.2
A:Map position: 5
A:Introns: 175/1; 209/3; 304/2; 482/3; 611/3; 680/1; 787/1; 1048/3; 1112/2; 1197/2;
C:Superfamily: Caenorhabditis elegans hypothetical protein C14B4.2

Query Match 90.0%; Score 36; DB 2; Length 1404;
Best Local Similarity 83.3%; Pred. No. 87;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
||:|
Db - 843 CHVDC 848

RESULT 11

```
156985
kalinin B1 - mouse
C:Species: Mus sp. (mouse)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 13-Aug-1999
C:Accession: I56985
R:Utani, A.; Kopp, J.B.; Kozak, C.A.; Matsuki, Y.; Amizuka, N.; Sugiyama, S.; Yamada, Y.
Lab. Invest. 72, 300-310, 1995
A>Title: Mouse kalinin B1 (laminin beta 3 chain): cloning and tissue distribution.
A:Reference number: I56985; MUID:95205823; PMID:7898049
A:Accession: I56985
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1168 <RES>
A:CROSS-references: GB:S75486; NID:9912768; PID:9912769
C:Superfamily: laminin-type EGF-like homology
F:376-425/Domain: laminin-type EGF-like homology <LEG>
F:428-475/Domain: laminin-type EGF-like homology <EG11>
F:478-528/Domain: laminin-type EGF-like homology <LEG1>
F:531-575/Domain: laminin-type EGF-like homology <LEG8>

Query Match      87.5%; Score 35; DB 2; Length 1168;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
DB 425 CHACDC 410

RESULT 12
laminin beta-2 chain precursor (version 1) - human
A:Species: Homo sapiens (man)
C>Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 17-Mar-1999
C:Accession: A55677
R:Wewer, U.M.; Gerecke, D.R.; Durkin, M.E.; Kurtz, K.S.; Mattei, M.G.; Champlaud, M.P.;
Genomics 24, 243-252, 1994
A>Title: Human beta2 chain of laminin (formerly S chain): cDNA cloning, chromosomal loca
A:Reference number: A55677; MUID:95213013; PMID:7698745
A:Accession: A55677
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1797 <WEW>
A:CROSS-references: GB:X79683
C:Genetics:
A:Gene: GDB:LAMB2
A:CROSS-references: GDB:132363; OMIM:150325
A:Map position: 3p21.3-3p21.2
C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
C:Keywords: basement membrane; extracellular matrix; glycoprotein; heterotrimer
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-1797/Product: laminin beta-2 chain #status predicted <MAT>
F:283-344/Domain: laminin-type EGF-like homology <LE01>
F:347-407/Domain: laminin-type EGF-like homology <LE02>
F:410-467/Domain: laminin-type EGF-like homology <LE03>
F:470-519/Domain: laminin-type EGF-like homology <LE04>
F:522-552/Domain: laminin-type EGF-like homology #status atypical <LE05>
F:583-828/Domain: laminin-type EGF-like homology <LE06>
F:831-874/Domain: laminin-type EGF-like homology <LE07>
F:877-924/Domain: laminin-type EGF-like homology <LE08>
F:927-983/Domain: laminin-type EGF-like homology <LE09>
F:986-1034/Domain: laminin-type EGF-like homology <LE10>
F:1037-1091/Domain: laminin-type EGF-like homology <LE11>
F:1094-1139/Domain: laminin-type EGF-like homology <LE12>
F:1142-1196/Domain: laminin-type EGF-like homology <LE13>

Query Match      87.5%; Score 35; DB 2; Length 1797;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
DB 1139 CHACDC 1144
```

```
RESULT 13
S53869
laminin beta-2 chain precursor (version 2) - human
N:Alternate names: s-laminin
C:Species: Homo sapiens (man)
C>Date: 27-Oct-1995 #sequence_revision 23-Feb-1996 #text_change 24-Sep-1998
C:Accession: S53869
R:Jivanainen, A.; Vuolteenaho, R.; Sainio, K.; Eddy, R.; Shows, T.B.; Sariola, H.; T
Matrix Biol. 14, 489-497, 1994
A>Title: The human laminin beta-2 chain (S-laminin): structure, expression in fetal
A:Reference number: S53869
A:Accession: S53869
A:Molecule type: mRNA
A:Residues: 1-1798 <IIV>
C:Genetics:
A:Gene: GDB:LAMB2
A:CROSS-references: GDB:132363; OMIM:150325
A:Map position: 3p21.3-3p21.2
C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
C:Keywords: basement membrane; extracellular matrix; glycoprotein; heterotrimer
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-1798/Product: laminin beta-2 chain #status predicted <MAT>
F:283-344/Domain: laminin-type EGF-like homology <LE01>
F:347-407/Domain: laminin-type EGF-like homology <LE02>
F:410-467/Domain: laminin-type EGF-like homology <LE03>
F:470-519/Domain: laminin-type EGF-like homology <LE04>
F:522-552/Domain: laminin-type EGF-like homology #status atypical <LE05>
F:583-828/Domain: laminin-type EGF-like homology <LE06>
F:831-874/Domain: laminin-type EGF-like homology <LE07>
F:877-924/Domain: laminin-type EGF-like homology <LE08>
F:927-983/Domain: laminin-type EGF-like homology <LE09>
F:986-1034/Domain: laminin-type EGF-like homology <LE10>
F:1037-1091/Domain: laminin-type EGF-like homology <LE11>
F:1094-1140/Domain: laminin-type EGF-like homology <LE12>
F:1143-1187/Domain: laminin-type EGF-like homology <LE13>

Query Match      87.5%; Score 35; DB 2; Length 1798;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
DB 1140 CHACDC 1145

RESULT 14
AG2594
enoyl-(acyl-carrier-protein) reductase [NADH] fabI [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AG2594
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Wo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McC
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Ka
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG2594
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <KUR>
A:CROSS-references: GB:AE009688; PID:AA141173.1; PID:g17738472; GSPDB:GN00186
A:Experimental source: strain C58 (DuPont)
C:Genetics:
A:Gene: fabI
A:Superfamily: enoyl-(acyl-carrier-protein) reductase (NADH); short-chain alcohol de
Query Match      85.0%; Score 34; DB 2; Length 268;
```

Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 244 CHYVDC 249

RESULT 15

H97376
enoyl-(acyl-carrier-protein) reductase (NADH) (NADH-dependent enoyl-ACP reductase). [impo
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C/Accession: H97376
A: Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; PMID:11743194
A/Accession: H97376
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-270 <KUR>
A/Cross-references: GB:AE007869; PIDN:AK85969.1; PID:g15155028; GSPDB:GN00169
C/Genetics:
A/Map position: circular chromosome
A/Map position: circular chromosome
C/Superfamily: enoyl-(acyl-carrier-protein) reductase (NADH); short-chain alcohol dehydro

Query Match 85.0%; Score 34; DB 2; Length 270;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||
Db 246 CHYVDC 251

RESULT 16

A56609
uroporphyrinogen decarboxylase (EC 4.1.1.37) - Synecococcus sp. (strain PCC 7942)
C/Species: Synecococcus sp.
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 31-Mar-2000
C/Accession: A56609; PQ0096; A30488
R/Kiel, J.A.; Ten Berge, A.M.; Venema, G.
DNA Seq. 2, 415-418, 1992
A/Title: Nucleotide sequence of the Synecococcus sp. PCC7942 heme gene encoding the heme
A/Reference number: A56609; MUID:93076004; PMID:1339332
A/Accession: A56609
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-354 <K12>
A/Cross-references: EMBL:Z11705; NID:g48039; PIDN:CAA77766.1; PID:g48040
A/Note: sequence extracted from NCBI backbone (NCBIP:118919)
R/Kiel, J.A.K.W.; Boels, J.M.; Beidman, G.; Venema, G.
Gene 89, 77-84, 1990
A/Title: Nucleotide sequence of the Synecococcus sp. PCC7942 branching enzyme gene (gls
A/Reference number: JQ0550; MUID:90323609; PMID:2142668
A/Accession: PQ0096
A/Molecule type: DNA
A/Residues: 1-79 <K1B>
A/Cross-references: GB:M31544; NID:g142134; PIDN:AAB39039.1; PID:g142136
C/Genetics:
A/Gene: heme
C/Superfamily: uroporphyrinogen decarboxylase
C/Keywords: carbon-carbon lyase; carboxy-lyase; porphyrin biosynthesis

Query Match 85.0%; Score 34; DB 2; Length 354;
Best Local Similarity 66.7%; Pred. No. 64;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||

Db 195 CHQIDC 200

RESULT 17

D83305
hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PAO1)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: D83305
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; I

Query Match 82.5%; Score 33; DB 2; Length 130;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||

RESULT 18

D86236
protein F14N23.5 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: D86236
R/theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
anzen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallic
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: D86236
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1188 <STO>
A/Cross-references: GB:AE005172; NID:g4914319; PIDN:AAD32867.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1

Query Match 83.3%; Score 33; DB 2; Length 1188;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||

RESULT 19

D95961
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid ps
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C;Accession: D95961
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D95961
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-152 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49356.1; PID:g15140842; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymb
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMD21378
A;Genome: plasmid

Query Match 80.0%; Score 32; DB 2; Length 152;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 54 CHCADC 59

RESULT 20

S31829
pathogenesis-related protein P23 precursor - tomato (fragment)
C;Species: Lycopersicon esculentum (tomato)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
C;Accession: PQ0742; S31829
R;Rodrigo, I.; Vera, P.; Tornero, P.; Hernandez-Yago, J.; Conejero, V.
Plant Physiol. 102, 939-945, 1993
A;Title: cDNA cloning of viroid-induced tomato pathogenesis-related protein P23. Charac
A;Reference number: PQ0742; MUID:94105335; PMID:8278538
A;Accession: PQ0742
A;Molecule type: mRNA
A;Residues: 1-233 <RO2>
A;Cross-references: EMBL:X70787; NID:g19314; PIDN:CAA50059.1; PID:g19315
A;Experimental source: leaf, cv. Rutgers, cDNA:PTCP23.1
C;Superfamily: thaumatin I
F;1-8/Domain: signal sequence (fragment) #status predicted <SIG>
F;9-233/Product: pathogenesis-related protein P23 #status predicted <MAT>

Query Match 80.0%; Score 32; DB 2; Length 233;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 128 CHAHC 133

RESULT 21

S28001
osmotin-like protein TPm1 precursor - tomato (fragment)
C;Species: Lycopersicon esculentum (tomato)
C;Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 20-Aug-1999
C;Accession: S28001; S22073
R;Ruiz-Medrano, R.; Jimenez-Morilla, B.; Herrera-Estrella, L.; Rivera-Bustamante, R.F.
Plant Mol. Biol. 20, 1199-1202, 1992
A;Title: Nucleotide sequence of an osmotin-like cDNA induced in tomato during viroid inf
A;Reference number: S28001; MUID:93099273; PMID:1463856
A;Accession: S28001
A;Molecule type: mRNA

A;Residues: 1-238 <RUI>
A;Cross-references: EMBL:X66416; NID:g22628; PIDN:CAA47047.1; PID:g22629
C;Genetics:
A;Gene: TPm1
C;Superfamily: thaumatin I
F;1-13/Domain: signal sequence (fragment) #status predicted <SIG>
F;14-238/Product: osmotin-like protein TPm1 #status predicted <MAT>

Query Match 80.0%; Score 32; DB 2; Length 238;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 133 CHAHC 138

RESULT 22

S07406
thaumatin homolog NP24 precursor - tomato (fragment)
N;Alternate names: salt-induced protein NP24
C;Species: Lycopersicon esculentum (tomato)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 20-Aug-1999
C;Accession: S07406; PS0283; JAO177
R;King, G.J.; Turner, V.A.; Hussey Jr., C.E.; Wurtele, E.S.; Lee, S.M.
Plant Mol. Biol. 10, 401-412, 1988
A;Title: Isolation and characterization of a tomato cDNA clone which codes for a
A;Reference number: S07406
A;Accession: S07406
A;Molecule type: mRNA
A;Residues: 1-239 <KINI>
A;Cross-references: EMBL:M21346; NID:g170466; PIDN:AAA34175.1; PID:g170467
A;Experimental source: suspension cells
A;Note: part of this sequence, including the amino end of the mature protein, was c
A;Accession: PS0283
A;Molecule type: protein
A;Residues: 14-93 <KIN2>
C;Comment: NP24 protein accumulates after cells are stressed with agents, such as N
C;Superfamily: thaumatin I
F;1-13/Domain: signal sequence (fragment) #status predicted <SIG>
F;14-239/Product: thaumatin homolog NP24 #status predicted <MAT>

Query Match 80.0%; Score 32; DB 2; Length 239;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 133 CHAHC 138

RESULT 23

S30144
osmotin-like protein precursor (clone pA13) - Commerson's wild potato
C;Species: Solanum commersonii (Commerson's wild potato)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Mar-2001
C;Accession: S30144; S33195; S25019
R;Zhu, B.; Chen, T.H.H.; Li, P.H.
Plant Mol. Biol. 21, 729-735, 1993
A;Title: Expression of an ABA-responsive osmotin-like gene during the induction of
A;Reference number: S30144; MUID:93192535; PMID:8448373
A;Accession: S30144
A;Molecule type: mRNA
A;Residues: 1-246 <ZHU>
A;Cross-references: EMBL:X67121; NID:g71194; PIDN:CAA47601.1; PID:g21195
A;Experimental source: clone pA13
R;Zhu, B.; Chen, T.H.H.; Li, P.H.
submitted to the EMBL Data Library, March 1993
A;Description: Nucleotide sequences of two ABA-responsive genes in potato.

A;Reference number: S33196
A;Accession: S33195
A;Molecule type: DNA
A;Residues: 1-246 <ZHW>

A;Cross-references: EMBL:X72928; NID:g296771; PIDN:CAA51432.1; PID:g296772
C;Superfamily: thaumatin I
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-246/Product: osmotin-like protein #status predicted <MAT>

Query Match 80.0%; Score 32; DB 2; Length 246;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 141 CHAHC 146

RESULT 24

S33197

osmotin-like protein precursor (clone pAB1) - Commerson's wild potato
C;Species: Solanum commersonii (Commerson's wild potato)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
C;Accession: S33197; S56709
R;Zhu, B.; Chen, T.H.H.; Li, P.H.
submitted to the EMBL Data Library, March 1993
A;Description: Isolation of cDNAs encoding osmotin-like proteins in potato.
A;Reference number: S33195
A;Accession: S33197
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-247 <ZHU>
A;Cross-references: EMBL:X72926; NID:g296775; PIDN:CAA51430.1; PID:g296776
R;Zhu, B.; Chen, T.H.H.; Li, P.H.
Plant Mol. Biol. 28, 17-26, 1995
A;Title: Expression of three osmotin-like protein genes in response to osmotic stress and
A;Reference number: S56708; MUID:95306785; PMID:7787181
A;Accession: S56709
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-247 <ZH>
A;Cross-references: EMBL:X72926; NID:g296775; PIDN:CAA51430.1; PID:g296776
C;Superfamily: thaumatin I

Query Match 80.0%; Score 32; DB 2; Length 247;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 141 CHAHC 146

RESULT 25

S33196

osmotin-like protein - Commerson's wild potato
C;Species: Solanum commersonii (Commerson's wild potato)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C;Accession: S33196
R;Zhu, B.; Chen, T.H.H.; Li, P.H.
submitted to the EMBL Data Library, March 1993
A;Description: Nucleotide sequences of two ABA-responsive genes in potato.
A;Reference number: S33196
A;Accession: S33196
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-247 <ZHU>
A;Cross-references: EMBL:X72927; NID:g296773; PIDN:CAA51431.1; PID:g296774
C;Superfamily: thaumatin I

Query Match 80.0%; Score 32; DB 2; Length 247;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
Db 141 CHAHC 146

Search completed: June 20, 2003, 20:59:15
Job time : 24.2727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 ; Search time 7.22727 Seconds
(without alignments)
34.433 Million cell updates/sec

Title: US-10-105-008-20
Perfect score: 40
Sequence: 1 CHAVDC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	92.5	225	UL04_HSVB	P28943 equine herp
2	36	90.0	169	X193_ECOLI	Q00739 escherichia
3	36	90.0	169	X19F_ECOLI	P47737 escherichia
4	36	90.0	712	FBL1_CAEEL	O77469 caenorhabdi
5	35	87.5	1168	LMB3_MOUSE	Q61087 mus musculus
6	35	87.5	1798	LMB2_HUMAN	P55268 homo sapien
7	34	85.0	354	DCUP_SYNP7	P18891 synechococc
8	34	85.0	528	ORDA_ASPFL	P79084 aspergillus
9	33	82.5	142	PA2E_HUMAN	Q9nzk7 homd sapien
10	33	82.5	142	PA2E_MOUSE	Q9nzk7 homd sapien
11	33	82.5	168	PA2F_HUMAN	Q9bzm2 homo sapien
12	33	82.5	168	PA2F_MOUSE	Q9bzm2 homo sapien
13	33	82.5	528	ORDA_ASPFL	Q9nzk7 mus musculus
14	33	82.5	958	HIG_DROME	O13345 aspergillus
15	33	82.5	1829	Y296_HUMAN	Q09101 drosophila
16	32	80.0	238	TPW1_LYCES	O15015 homo sapien
17	32	80.0	246	OS13_LYCES	Q01591 lycopersico
18	32	80.0	246	OSMO_TOBAC	P50701 solanum com
19	32	80.0	247	NP24_LYCES	P14170 nicotiana t
20	32	80.0	247	OS81_SOLCO	P12670 lycopersico
21	32	80.0	549	LEM2_RAT	P98105 rattus norveg

34	31	77.5	668	1	VNCS_PAVCN
35	31	77.5	1053	1	SPS_SOLTU
36	31	77.5	1886	1	POL_COYMY
37	31	77.5	2048	1	RRPL_SENDE
38	31	77.5	2228	1	RRPL_SENDE
39	31	77.5	2228	1	RRPL_SENDF
40	31	77.5	2228	1	RRPL_SENDZ
41	31	77.5	3033	1	POLG_HCVJ6
42	30	75.0	82	1	C551_AZOVI
43	30	75.0	104	1	C551_PSEAE
44	30	75.0	311	1	ZITB_ECO57
45	30	75.0	312	1	ZITB_SALTI
46	30	75.0	312	1	ZITB_SALTY
47	30	75.0	313	1	ZITB_ECOLI
48	30	75.0	449	1	URPA_SCHPO
49	30	75.0	458	1	RADA_HAEIN
50	30	75.0	519	1	PDAS_HUMAN
51	30	75.0	668	1	VNCS_PAVL3
52	30	75.0	672	1	VNCS_MUMIV
53	30	75.0	672	1	VNCS_MUMIV
54	30	75.0	672	1	VNCS_PAVHH
55	30	75.0	1056	1	SPS_SPIOL
56	30	75.0	1134	1	VGLM_HANTH
57	30	75.0	1135	1	VGLM_HANTL
58	30	75.0	1135	1	VGLM_HANTV
59	30	75.0	1172	1	LMB3_HUMAN
60	30	75.0	1214	1	BRF3_HUMAN
61	30	75.0	1959	1	AGRI_RAT
62	30	75.0	2233	1	RRPL_P13H4
63	29	72.5	59	1	RL32_CHLPN
64	29	72.5	155	1	PA2X_HUMAN
65	29	72.5	169	1	Y224_METJA
66	29	72.5	172	1	TASM_POVBA
67	29	72.5	172	1	TASM_POVBK
68	29	72.5	173	1	POF5_YEAST
69	29	72.5	194	1	TASM_POVHA
70	29	72.5	251	1	OLPA_TOBAC
71	29	72.5	262	1	VER1_ASPPA
72	29	72.5	401	1	TAM1_POVHA
73	29	72.5	411	1	MSK1_MEDSA
74	29	72.5	560	1	HCVG_SEPOF
75	29	72.5	606	1	SP2_HUMAN

ALIGNMENTS

RESULT 1				
UL04_HSVB	UL04_HSVB	STANDARD;	PRT;	225 AA.
AC	P28943			
DT	01-DEC-1992 (Rel. 24, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DE	01-DEC-1992 (Rel. 24, Last annotation update)			
GN	Gene 58 protein.			
OS	Equine herpesvirus type 1 (strain Abap) (EHV-1).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Varicellovirus.			
OX	NCBI_TaxID=31520;			
DN				

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR ENBL; M86664; AB02493.1; --
DR PIR; D36801; WZBEF1.
DR InterPro; IPR004958; Herpes UL4.
DR Pfam; PF03277; Herpes UL4; 1.
SQ SEQUENCE 225 AA; 24287 MW; 870438EC6A20DB6A CRC64;

Query Match 92.58; Score 37; DB 1; Length 225;
Best Local Similarity 83.3%; Pred. No. 3.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:|
Db 149 CHALDC 154

RESULT 2
X193 ECOLI STANDARD; PRT; 169 AA.
AC Q00739; 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE X polypeptide (ORF 19) (P19 protein) (ORF169).
GN X OR 19
OS Escherichia coli.
OG Plasmid IncFI P307.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=711;
RX MEDLINE=91261994; PubMed=2096398;
RA Graus-Goldner A., Graus H., Schlacher T., Hoegenauer G.;
RT "The sequences of genes bordering orit in the enterotoxin plasmid
P307: comparison with the sequences of plasmids F and R1.";
RL Plasmid 24:119-131(1990).
CC -1- SIMILARITY: BELONGS TO THE IAGB/IPGF/P19 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation--
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR ENBL; M62986; AAA25521.1; --
DR PIR; A37390; A37390.
DR InterPro; IPR000189; SLT_domain.
DR Pfam; PF01464; SLT; 1.
KW Plasmid.
SQ SEQUENCE 169 AA; 19327 MW; F9123E8DCD9BCBAA CRC64;

Query Match 90.0%; Score 36; DB 1; Length 169;
Best Local Similarity 83.3%; Pred. No. 4.4;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:|
Db 17 CHATDC 22

RESULT 3
X19F ECOLI STANDARD; PRT; 169 AA.
AC P47737; 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE X polypeptide (ORF 19) (P19 protein) (ORF169).
GN YUBQ OR YGPA OR X OR 19.
OS Escherichia coli.
OG Plasmid F.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90136505; PubMed=2693941;
RA Itoh S., Cram D., Skurray R.;
RT "Nucleotide sequence of the leading region adjacent to the origin of
RT transfer on plasmid F and its conservation among conjugative
RT plasmids.";
RL Mol. Gen. Genet. 219:177-186(1989).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=X12 / CR63;
RA Shimizu H., Saitoh Y., Suda Y., Uehara K., Sampei G., Mizobuchi K.;
RT "Complete nucleotide sequence of the F plasmid: its implications for
RT organization and diversification of plasmid genomes.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-15 FROM N.A.
RX MEDLINE=84236116; PubMed=6329741;
RA Thompson R., Taylor L., Kelly K., Everett R., Willetts N.;
RT "The F plasmid origin of transfer: DNA sequence of wild-type and
RT mutant origins and location of origin-specific nicks.";
RL EMBO J. 3:1175-1180(1984).
CC -1- SIMILARITY: BELONGS TO THE IAGB/IPGF/P19 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation--
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR ENBL; X17539; CAA35576.1; --
DR ENBL; AF106329; AAA99219.1; --
DR ENBL; AF001918; BAA97940.1; --
DR ENBL; X00545; CAA25215.1; --
DR EcoGene; EG40084; YUBQ.
DR InterPro; IPR000189; SLT_domain.
DR Pfam; PF01464; SLT; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 169 AA; 19285 MW; EED73EDA80818BD9 CRC64;

Query Match 90.0%; Score 36; DB 1; Length 169;
Best Local Similarity 83.3%; Pred. No. 4.4;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:|
Db 17 CHATDC 22

RESULT 4
FBL1 CAEEL STANDARD; PRT; 712 AA.
ID FBL1 CAEEL
AC O77469; O77474; Q95NZ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibulin-1 precursor.
GN FBLN1 OR F56H11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Feloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=99120531; PubMed=9923656;
RA Barth J.L., Agraves K.M., Roark E.F., Little C.D., Agraves W.S.;
RT "Identification of chicken and C. elegans fibulin-1 homologs and
RL characterization of the C. elegans fibulin-1 gene.";
RN Matrix Biol. 17:635-646(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lloyd C.R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX (BY SIMILARITY).
CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; C (SHOWN HERE) AND D; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -|- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF051403; AAC28323.1; -;
DR EMBL; AF051403; AAC28324.1; -;
DR EMBL; AF051401; AAC28321.1; -;
DR EMBL; AF051402; AAC28322.1; -;
DR EMBL; Z68219; CAC35826.1; -;
DR EMBL; Z68749; CAC35826.1; JOINED.
DR EMBL; Z68219; CAC35827.1; -;
DR EMBL; Z68749; CAC35827.1; JOINED.
DR EMBL; Z68749; CAC35817.1; -;
DR EMBL; Z68219; CAC35817.1; JOINED.
DR EMBL; Z68749; CAC35818.1; -;
DR EMBL; Z68219; CAC35818.1; JOINED.
DR WormPep; F56H11.1a; CE26701.
DR WormPep; F56H11.1b; CE26702.
DR HSP; F16109; IFSB.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001891; EGF_Ca.
DR Pfam; PF00008; EGF; 6.
DR Pfam; PF01821; ANATO; 2.
DR SMART; SM00104; ANATO; 2.
DR SMART; SM00179; EGF_Ca; 6.
DR SMART; SM00001; EGF_like; 3.
DR PROSITE; PS00010; ASX HYDROXYL; 4.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_Ca; 8.
KW Signal; Glycoprotein; Extracellular matrix; Repeat; EGF-like domain;
KW Calcium-binding; Alternative splicing.
FT SIGNAL 1 17 POTENTIAL.

FT DISULFID 79 95 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 98 128 BY SIMILARITY.
FT DISULFID 111 129 BY SIMILARITY.
FT DISULFID 159 168 BY SIMILARITY.
FT DISULFID 164 178 BY SIMILARITY.
FT DISULFID 180 193 BY SIMILARITY.
FT DISULFID 199 212 BY SIMILARITY.
FT DISULFID 206 221 BY SIMILARITY.
FT DISULFID 227 240 BY SIMILARITY.
FT DISULFID 246 259 BY SIMILARITY.
FT DISULFID 253 268 BY SIMILARITY.
FT DISULFID 291 302 BY SIMILARITY.
FT DISULFID 308 320 BY SIMILARITY.
FT DISULFID 314 329 BY SIMILARITY.
FT DISULFID 336 349 BY SIMILARITY.
FT DISULFID 355 365 BY SIMILARITY.
FT DISULFID 360 374 BY SIMILARITY.
FT DISULFID 376 389 BY SIMILARITY.
FT DISULFID 395 409 BY SIMILARITY.
FT DISULFID 403 418 BY SIMILARITY.
FT DISULFID 420 433 BY SIMILARITY.
FT DISULFID 439 450 BY SIMILARITY.
FT DISULFID 446 459 BY SIMILARITY.
FT DISULFID 461 474 BY SIMILARITY.
FT DISULFID 480 495 BY SIMILARITY.
FT DISULFID 491 504 BY SIMILARITY.
FT DISULFID 506 519 BY SIMILARITY.
FT DISULFID 525 537 BY SIMILARITY.
FT DISULFID 530 546 BY SIMILARITY.
FT DISULFID 551 565 BY SIMILARITY.
FT VARSPLOC 564 712
RCNOPSACGLPECSKVPLFLTYQFISLARAVPISSHRPA
ITLKVSNAPNADTEVNPELOLKTIVGAPNVLPAIRANFL
LQGEKNSAVTLRSDLDGFTVLQQLLRMSKKGKGNFT
YAANLIVDAVAAKHNTVHPPLMKIR -> QIADGYSCKV
CSTEDTECLNGHTVLYQFRAVPSLKTIIISPIEVSRIVTH
MGVPFSDYNDLYVQGRHFRIVQERNIGIVOLVKPISGPTV
ETIKVNIHTKSRTGVILAFNEAITEISVSKYFP (IN
ISOFORM D)
RCNOPSACGLPECSKVPLFLTYQFISLARAVPISSHRPA
ITLKVSNAPNADTEVNPELOLKTIVGAPNVLPAIRANFL
LQGEKNSAVTLRSDLDGFTVLQQLLRMSKKGKGNFT
YAANLIVDAVAAKHNTVHPPLMKIR -> QIADGYSCKV
CSTEDTECLNGHTVLYQFRAVPSLKTIIISPIEVSRIVTH
MGVPFSDYNDLYVQGRHFRIVQERNIGIVOLVKPISGPTV
ETIKVNIHTKSRTGVILAFNEAITEISVSKYFP (IN
ISOFORM D)
SQ SEQUENCE 712 AA; 77009 MW; 52CE8CF8BF296BC5 CRC64;
Query Match 90.0%; Score 36; DB 1; Length 712;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 CHAVDC 6
DB 546 CHSIDC 551
RESULT 5
LMB3_MOUSE
ID LMB3_MOUSE STANDARD; PRT; 1168 AA.
AC Q61087;
DT 01-NOV-1997 /dal 16

CC is thought to mediate the attachment, migration, and organization
FT DISULFID 450 459 BY SIMILARITY.
FT DISULFID 462 475 BY SIMILARITY.

is thought to mediate the attachment, migration, and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components.

-1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three different polypeptide chains (alpha, beta, gamma), which are bound to each other by disulfide bonds into a cross-shaped molecule comprising one long and three short arms with globules at each end.

THE BETA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/NICLIN).

-1- SUBCELLULAR LOCATION: Extracellular.

-1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR COMPONENT).

-1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.

-1- DOMAIN: DOMAIN VI IS GLOBULAR.

-1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).

-1- SIMILARITY: CONTAINS 6 LAMININ EGF-LIKE DOMAINS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U43298; AAA85255.1; -
HSSP: P02468; 1KLO.
MGD: MGI:99915; Lamb3.
InterPro: IPR000561; EGF-like.
InterPro: IPR001886; LamNT.
InterPro: IPR002049; Laminin EGF.
Pfam: PF00053; laminin EGF; 6.
Pfam: PF00055; laminin Nterm; 1.
PRINTS: PR00011; EGF/LAMININ.
ProDom: PD002082; LamNT; 1.
SMART: SM00180; EGF Lam; 6.
SMART: SM00136; LamNT; 1.
PROSITE: PS00022; EGF 1; 4.
PROSITE: PS01186; EGF 2; 2.
PROSITE: PS01248; LAMININ TYPE EGF; 3.
Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
Laminin EGF-like domain; Cell adhesion; Repeat; Signal.

FT CHAIN 1 17
FT SIGNAL 18 17
FT DOMAIN 18 1168 LAMININ BETA-3 CHAIN.
FT DOMAIN 18 249 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 250 312 LAMININ EGF-LIKE 1.
FT DOMAIN 313 375 LAMININ EGF-LIKE 2.
FT DOMAIN 376 427 LAMININ EGF-LIKE 3.
FT DOMAIN 428 477 LAMININ EGF-LIKE 4.
FT DOMAIN 478 530 LAMININ EGF-LIKE 5.
FT DOMAIN 531 575 LAMININ EGF-LIKE 6.
FT DOMAIN 576 781 DOMAIN II.
FT DOMAIN 782 812 DOMAIN ALPHA.
FT DOMAIN 813 1168 DOMAIN I.
FT DOMAIN 732 754 COILED COIL (POTENTIAL).
FT DOMAIN 827 879 COILED COIL (POTENTIAL).
FT DOMAIN 944 1129 COILED COIL (POTENTIAL).
FT DISULFID 250 259 BY SIMILARITY.
FT DISULFID 252 276 BY SIMILARITY.
FT DISULFID 278 287 BY SIMILARITY.
FT DISULFID 290 310 BY SIMILARITY.
FT DISULFID 313 322 BY SIMILARITY.
FT DISULFID 315 340 BY SIMILARITY.
FT DISULFID 343 352 BY SIMILARITY.
FT DISULFID 355 373 BY SIMILARITY.
FT DISULFID 376 389 BY SIMILARITY.
FT DISULFID 378 396 BY SIMILARITY.
FT DISULFID 398 407 BY SIMILARITY.
FT DISULFID 410 425 BY SIMILARITY.
FT DISULFID 428 441 BY SIMILARITY.
FT DISULFID 430 448 BY SIMILARITY.

FT DISULFID 450 459 BY SIMILARITY.
FT DISULFID 462 475 BY SIMILARITY.
FT DISULFID 480 490 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 516 528 BY SIMILARITY.
FT DISULFID 531 543 BY SIMILARITY.
FT DISULFID 550 550 BY SIMILARITY.
FT DISULFID 553 561 BY SIMILARITY.
FT DISULFID 564 575 BY SIMILARITY.
FT DISULFID 578 578 INTERCHAIN (PROBABLE).
FT DISULFID 581 581 INTERCHAIN (PROBABLE).
FT DISULFID 1167 1167 INTERCHAIN (PROBABLE).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 806 806 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1168 AA; 128996 MW; 31E22FA8ADD6EBC0 CRC64;

Query Match 87.5%; Score 35; DB 1; Length 1168;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CHAVDC 6
DB 425 CHACDC 430

RESULT 6
LMB2 HUMAN
ID LMB2 HUMAN STANDARD; PRT: 1798 AA.
AC P55268; Q16321;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin beta-2 chain precursor (S-laminin) (Laminin Bis chain).
GN LAMB2 OR LAMS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213013; PubMed=7698745;
RA Wewer U.M., Gerecke D.R., Durkin M.E., Kurtz K.S., Mattei M.-G., Champelaud M.F., Burgeson R.E., Albrechtsen R.;
RT "Human beta 2 chain of laminin (formerly S chain): cDNA cloning, chromosomal localization, and expression in carcinomas.";
RL Genomics 24:243-252(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95316263; PubMed=7795887;
RA Iivanainen A., Vuolteenaho R., Sainio K., Eddy R., Shows T.B., Sariola H., Tryggvason K.;
RT "The human laminin beta 2 chain (S-laminin): structure, expression in fetal tissues and chromosomal assignment of the LAMB2 gene.";
RL Matrix Biol. 14:489-497(1995).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin is thought to mediate the attachment, migration, and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three different polypeptide chains (alpha, beta, gamma), which are bound to each other by disulfide bonds into a cross-shaped molecule comprising one long and three short arms with globules at each end.
CC THE BETA-2 CHAIN IS A SUBUNIT OF LAMININ-3 (S-LAMININ), LAMININ-4 (S-MEROSIN), AND LAMININ-7 (KS-LAMININ).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR COMPONENT). S-LAMININ IS CONCENTRATED IN THE SYNAPTIC CLEFT OF THE NEUROMUSCULAR JUNCTION.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.

RA Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT "Nucleotide sequence of the *Synechococcus* sp. PCC7942 branching
RT enzyme gene (gigB): expression in *Bacillus subtilis*.";
RL Gene 89:77-84(1990).
CC -1- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4
CC CO(2).
CC -1- PATHWAY: Porphyrin biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z11705; CAAT7766.1; -;
DR EMBL; M31544; AAB39039.1; -;
DR F1R; PQ0096; PQ0096.
DR HSSP; P06132; IURO.
DR InterPro; IPR000257; Uro-decarboxyls.
DR Pfam; PF01208; URO-D; 1.
DR PROSITE; PS00906; UROD_1; 1.
DR PROSITE; PS00907; UROD_2; 1.
KW Lyase; Decarboxylase; Porphyrin biosynthesis;
KW Chlorophyll biosynthesis.
SQ SEQUENCE 354 AA; 39278 MW; 3F6846658734C1E9 CRC64;

Query Match 85.0%; Score 34; DB 1; Length 354;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
DB 195 CHQIDC 200

RESULT 8
ORDA ASPFL
ID ORDA ASPFL STANDARD; PRT; 528 AA.
AC P79084;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE O-methylsterigmatocystin oxidoreductase (EC 1.14.1.-) (OMST
DE oxidoreductase) (Cytochrome P450 64).
GN ORDI OR CYP64.
OS *Aspergillus flavus*.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
OX NCBI_TaxID=5059;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 3357;
RX MEDLINE=97288058; PubMed=9143099;
RA Prieto R., Woloshuk C.P.;
RT "ordi, an oxidoreductase gene responsible for conversion of O-
RT methylsterigmatocystin to aflatoxin in *Aspergillus flavus*.";
RL Appl. Environ. Microbiol. 63:1661-1666(1997).
CC -1- FUNCTION: CONVERTS O-METHYLSTERIGMATOCYSTIN (OMST) TO AFLATOXIN
CC B1.
CC
CC -1- PATHWAY: Aflatoxin biosynthesis.
CC
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC EMBL; U81806; AAC49709.1; -;
DR EMBL; U81807; AAC49710.1; -;
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 2.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Heme.
FT BINDING 440 440 HEME (BY SIMILARITY).
SQ SEQUENCE 528 AA; 60210 MW; 56DEA320008BD5EE CRC64;

Query Match 85.0%; Score 34; DB 1; Length 528;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
DB 455 CHAVSC 460

RESULT 9
PA2E HUMAN
ID PA2E HUMAN STANDARD; PRT; 142 AA.
AC Q9NZK7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Group IIE secretory phospholipase A2 precursor (EC 3.1.1.4)
DE (Phosphatidylcholine 2-acylhydrolase GIIE) (GIIE SPLA2) (sPLA(2)-IIE).
GN PLA2G2E
OS *Homo sapiens* (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20148788; PubMed=10681567;
RA Suzuki N., Ishizaki J., Yokota Y., Higashino K., Ono T., Ikeda M.,
RA Fujii N., Kawamoto K., Hanasaki K.;
RT "Structures, enzymatic properties, and expression of novel human and
RT mouse secretory phospholipase A(2)s.";
RL J. Biol. Chem. 275:5785-5793(2000).
CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
CC acyl groups in 3-sn-phosphoglycerides. Has a preference for
CC arachidonic-containing phospholipids.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
CC acylglycerophosphocholine + a fatty acid anion.
CC -1- COFACTOR: Binds 1 calcium ion per subunit.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Restricted to the brain, heart, lung, and
CC placenta.
CC
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF189279; AAF36541.1; -;
DR HSSP; P14555; IPOD.
DR InterPro; IPR001211; PhospholipaseA2.
DR Pfam; PF00068; phoslip; 1.
DR PRINTS; P00389; PHPLIPASEA2.
DR PRODOM; PD000303; PhospholipaseA2; 1.
DR SMART; SM00085; PA2c; 1.
DR PROSITE; PS00119; PA2 ASP; FALSE_NEG.
DR PROSITE; PS00118; PA2 HIS; 1.
KW Hydrolase; Lipid degradation; Signal; Calcium.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 142 GROUP IIE SECRETORY PHOSPHOLIPASE A2.
FT ACT_SITE 65 65 BY SIMILARITY.

FT ACT_SITE 109 109 BY SIMILARITY.
 FT DISULFID 44 135 BY SIMILARITY.
 FT DISULFID 46 62 BY SIMILARITY.
 FT DISULFID 61 115 BY SIMILARITY.
 FT DISULFID 67 142 BY SIMILARITY.
 FT DISULFID 68 108 BY SIMILARITY.
 FT DISULFID 77 101 BY SIMILARITY.
 FT DISULFID 95 106 BY SIMILARITY.
 FT CA_BIND 45 45 VIA CARBONYL OXYGEN (BY SIMILARITY).
 FT CA_BIND 47 47 VIA CARBONYL OXYGEN (BY SIMILARITY).
 FT CA_BIND 49 49 VIA CARBONYL OXYGEN (BY SIMILARITY).
 FT CA_BIND 66 66 BY SIMILARITY.
 SQ SEQUENCE 142 AA; 15989 MW; 3C360EA710E141FB CRC64;

Query Match 82.5%; Score 33; DB 1; Length 142;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 62 CHAHDC 67

RESULT 10
 PAZE_MOUSE STANDARD; PRT; 142 AA.
 AC Q9UL3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Group IIE secretory phospholipase A2 precursor (EC 3.1.1.4)
 DE (Phosphatidylcholine 2-acylhydrolase GIIIE) (GIIIE sPLA2) (sPLA(2)-IIE).
 GN PLA2G2E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20002639; PubMed=10531313;
 RA Valentin E., Ghomashchi F., Gelb M.H., Lazdunski M., Lambeau G.;
 RT "On the diversity of secreted phospholipases A2. Cloning, tissue
 RT distribution, and functional expression of two novel mouse group II
 RT enzymes.";
 RL J. Biol. Chem. 274:31195-31202 (1999).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=BALE/c;
 RX MEDLINE=20148788; PubMed=10681567;
 RA Suzuki N., Ishizaki J., Yokota Y., Higashino K., Ono T., Ikeda M.,
 RA Fujii N., Kawamoto K., Hanasaki K.;
 RT "Structures, enzymatic properties, and expression of novel human and
 RT mouse secretory phospholipase A(2)s.";
 RL J. Biol. Chem. 275:5785-5793 (2000).
 CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
 CC acyl groups in 3-sn-phosphoglycerides.
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
 CC acylglycerophosphocholine + a fatty acid anion.
 CC -1- COFACTOR: Binds 1 calcium ion per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Highly expressed in uterus, and at lower
 CC levels in various other tissues.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC and this statement is not removed. Usage by and for commercial
 CC purposes requires a separate agreement. (See <http://www.isb-sib.ch/announce/>)

DR EMBL; AF112984; AAF22290.1; -.
 DR HSP; P14555; IPOD.
 DR MGD; MGI:1349660; Pla2g2e.
 DR InterPro; IPR001211; PhospholipaseA2.
 DR Pfam; PF00068; phoslip; 1.
 DR PRINTS; PR00389; PPHLIPASEA2.
 DR ProDom; PD00303; PhospholipaseA2; 1.
 DR SMART; SM00085; PA2c; 1.
 DR PROSITE; PS00119; PA2 ASP; FALSE_NEG.
 DR PROSITE; PS00118; PA2_HIS; 1.
 KW Hydrolase; Lipid degradation; Signal; Calcium.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 142 GROUP IIE SECRETORY PHOSPHOLIPASE A2.
 FT ACT_SITE 65 65 BY SIMILARITY.
 FT ACT_SITE 109 109 BY SIMILARITY.
 FT DISULFID 44 135 BY SIMILARITY.
 FT DISULFID 46 62 BY SIMILARITY.
 FT DISULFID 61 115 BY SIMILARITY.
 FT DISULFID 67 142 BY SIMILARITY.
 FT DISULFID 77 101 BY SIMILARITY.
 FT DISULFID 95 106 BY SIMILARITY.
 FT CA_BIND 45 45 VIA CARBONYL OXYGEN (BY SIMILARITY).
 FT CA_BIND 47 47 VIA CARBONYL OXYGEN (BY SIMILARITY).
 FT CA_BIND 49 49 VIA CARBONYL OXYGEN (BY SIMILARITY).
 FT CA_BIND 66 66 BY SIMILARITY.
 SQ SEQUENCE 142 AA; 15942 MW; 8B0E3CC710A1F946 CRC64;

Query Match 82.5%; Score 33; DB 1; Length 142;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 62 CHAHDC 67

RESULT 11
 PA2F_HUMAN STANDARD; PRT; 168 AA.
 ID Q9BZW2; Q9H506;
 AC Q9BZW2; Q9H506;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Group IIF secretory phospholipase A2 precursor (EC 3.1.1.4)
 DE (Phosphatidylcholine 2-acylhydrolase GIIF) (GIIF sPLA2) (sPLA(2)-IIF).
 GN PLA2G2F.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=20563827; PubMed=11112443;
 RA Valentin E., Singer A.G., Ghomashchi F., Lazdunski M., Gelb M.H.,
 RA Lambeau G.;
 RT "Cloning and recombinant expression of human group IIF-secreted
 RT phospholipase A(2).";
 RL Biochem. Biophys. Res. Commun. 279:223-228 (2000).
 RN [2]
 RP SEQUENCE OF 1-98 FROM N.A.
 RA Wallis J.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
 CC acyl groups in 3-sn-phosphoglycerides. Hydrolyzes
 CC phosphatidylglycerol versus phosphatidylcholine with a 15-fold
 CC preference.
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
 CC acylglycerophosphocholine + a fatty acid
 CC -1- COFACTOR: Binds 1 calcium ion per
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Spleen and

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:47:18 ; Search time 14.7955 Seconds
(without alignments)
13.921 Million cell updates/sec

Title: US-10-105-008-44
Perfect score: 42
Sequence: 1 CHAVSSC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	42	100.0	7	4	US-08-996-679-54
2	42	100.0	7	4	US-09-115-395-68
3	42	100.0	7	4	US-09-250-059-44
4	42	100.0	7	4	US-09-248-074-44
5	42	100.0	7	4	US-09-357-717-28
6	42	100.0	7	4	US-09-458-870-44
7	38	90.5	7	4	US-09-115-395-69
8	38	90.5	7	4	US-09-250-059-45
9	38	90.5	7	4	US-09-248-074-45
10	38	90.5	7	4	US-09-458-870-45
11	33	78.6	8	3	US-08-893-534A-18
12	33	78.6	8	4	US-08-996-679-18
13	33	78.6	8	4	US-09-115-395-43
14	33	78.6	8	4	US-09-507-102-18
15	33	78.6	8	4	US-09-250-059-42
16	33	78.6	8	4	US-09-248-074-42
17	33	78.6	8	4	US-09-357-717-27
18	33	78.6	8	4	US-09-458-870-42
19	33	78.6	9	3	US-08-893-534A-31
20	33	78.6	9	3	US-08-893-534A-32
21	33	78.6	9	4	US-08-996-679-31
22	33	78.6	9	4	US-08-996-679-32
23	33	78.6	9	4	US-09-115-395-47
24	33	78.6	9	4	US-09-115-395-48
25	33	78.6	9	4	US-09-507-102-31
26	33	78.6	9	4	US-09-507-102-32
27	33	78.6	9	4	US-09-250-059-71

Sequence 72, Appl
Sequence 71, Appl
Sequence 72, Appl
Sequence 50, Appl
Sequence 51, Appl
Sequence 71, Appl
Sequence 72, Appl
Sequence 30, Appl
Sequence 30, Appl
Sequence 46, Appl
Sequence 30, Appl
Sequence 70, Appl
Sequence 70, Appl
Sequence 49, Appl
Sequence 70, Appl
Patent No. 5177197
Sequence 59, Appl
Sequence 179, Appl
Sequence 8, Appl
Sequence 50, Appl
Sequence 50, Appl
Sequence 31, Appl
Sequence 50, Appl
Sequence 7, Appl
Sequence 2, Appl
Sequence 11, Appl
Sequence 5, Appl
Sequence 2, Appl
Sequence 14, Appl
Sequence 15, Appl
Sequence 11, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 92, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 42, Appl
Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-996-679-54
; Sequence 54, Application US/08996679
; Patent No. 6169071
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

US-RATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: Blaschuk, Orest W.
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-54

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 1 CHAVSSC 7

RESULT 2

US-09-115-395-68
Sequence 68, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115,395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996,679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893,534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 68
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis
US-09-115-395-68

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 1 CHAVSSC 7

RESULT 3

US-09-250-059-44
Sequence 44, Application US/09250059
Patent No. 633307
GENERAL INFORMATION:

APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250,059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 44
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-250-059-44

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 1 CHAVSSC 7

RESULT 4

US-09-248-074-44
Sequence 44, Application US/09248074
Patent No. 6346512
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
FILE REFERENCE: 100086.401C5
CURRENT APPLICATION NUMBER: US/09/248,074
CURRENT FILING DATE: 1999-02-10
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 44
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-248-074-44

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 1 CHAVSSC 7

RESULT 5

US-09-357-717-28
Sequence 28, Application US/09357717
Patent No. 6417325

GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-357-717-28

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
DB 1 CHAVSSC 7

RESULT 6
US-09-458-870-44
; Sequence 44, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-44

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
DB 1 CHAVSSC 7

RESULT 7
US-09-115-395-69
; Sequence 69, Application US/09115395A
; Patent No. 6207639
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C3
; CURRENT APPLICATION NUMBER: US/09/115,395A
; CURRENT FILING DATE: 1998-07-14
; EARLIER APPLICATION NUMBER: 08/996,679
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/893,534
; EARLIER FILING DATE: 1997-07-11
; EARLIER APPLICATION NUMBER: 60/021,612
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
; OTHER INFORMATION: Synthesis
US-09-115-395-69

Query Match 90.5%; Score 38; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
DB 1 CHGVSSC 7

RESULT 8
US-09-250-059-45
; Sequence 45, Application US/09250059
; Patent No. 6333307
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
; FILE REFERENCE: 100086.401C6
; CURRENT APPLICATION NUMBER: US/09/250,059
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: control peptide
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-250-059-45

Query Match 90.5%; Score 38; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
DB 1 CHGVSSC 7

RESULT 9

US-09-248-074-45
; Sequence 45, Application US/09248074
; Patent No. 6346512
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C5
; CURRENT APPLICATION NUMBER: US/09/248,074
; CURRENT FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: control peptide
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-248-074-45

Query Match 90.5%; Score 38; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
DB 1 CHGVSSC 7

RESULT 10

US-09-458-870-45
; Sequence 45, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Gour, Barbara J.
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: control peptide
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-45

Query Match 90.5%; Score 38; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
DB 1 CHGVSSC 7

RESULT 11

US-08-893-534A-18
; Sequence 18, Application US/08893534A
; Patent No. 6031072
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,534A
; FILING DATE: 11-JUL-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 100086.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-893-534A-18

Query Match 78.6%; Score 33; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
|||
DB 3 HAVSSC 8

RESULT 12

US-08-996-679-18
; Sequence 18, Application US/08996679
; Patent No. 6169071
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
US-08-996-679-18

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
|||||
DB 3 HAVSSC 8

RESULT 13

US-09-115-395-43
Sequence 43, Application US/09115395A
Patent No. 6207639
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C3
CURRENT APPLICATION NUMBER: US/09/115,395A
CURRENT FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 08/996,679
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/893,534
EARLIER FILING DATE: 1997-07-11
EARLIER APPLICATION NUMBER: 60/021,612
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Solid Phase
OTHER INFORMATION: Synthesis
US-09-115-395-43

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
|||||
DB 3 HAVSSC 8

RESULT 14

US-09-507-102-18
Sequence 18, Application US/09507102
Patent No. 6326352
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.

Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED IP LAW GROUP PLLC
STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,102
FILING DATE: 17-Feb-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/893,534
FILING DATE: 11-JUL-1997
APPLICATION NUMBER: US 60/021,612
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Christiansen, William T.
REGISTRATION NUMBER: 44,614
REFERENCE/DOCKET NUMBER: 100086.401C10
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-507-102-18

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
|||||
DB 3 HAVSSC 8

RESULT 15

US-09-250-059-42
Sequence 42, Application US/09250059
Patent No. 6333307
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH
FILE REFERENCE: 100086.401C6
CURRENT APPLICATION NUMBER: US/09/250,059
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
FEATURE:

OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-250-059-42

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
| | | | |
Db 3 HAVSSC 8

RESULT 16
US-09-248-074-42
; Sequence 42, Application US/09248074
; Patent No. 6346512
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C5
; CURRENT APPLICATION NUMBER: US/09/248,074
; CURRENT FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-248-074-42

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
| | | | |
Db 3 HAVSSC 8

RESULT 17
US-09-357-717-27
; Sequence 27, Application US/09357717
; Patent No. 6417325
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C7
; CURRENT APPLICATION NUMBER: US/09/357,717
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion

OTHER INFORMATION: recognition sequence

FEATURE:
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
OTHER INFORMATION: and/or C-terminal modifications such as amide or
OTHER INFORMATION: ester group
US-09-357-717-27

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
| | | | |
Db 3 HAVSSC 8

RESULT 18
US-09-458-870-42
; Sequence 42, Application US/09458870
; Patent No. 6465427
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Farookhi, Riaz
; APPLICANT: Ali, Ammar
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CELL ADHESION
; FILE REFERENCE: 100086.401C8
; CURRENT APPLICATION NUMBER: US/09/458,870
; CURRENT FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; OTHER INFORMATION: ester group
US-09-458-870-42

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
| | | | |
Db 3 HAVSSC 8

RESULT 19
US-08-893-534A-31
; Sequence 31, Application US/08893534A
; Patent No. 6031072
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEED and BERRY LLP
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,534A
FILING DATE: 11-JUL-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /product= "OTHER"
OTHER INFORMATION: /note= "Residue is
US-08-893-534A-31

Query Match 78.6%; Score 33; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
Db 4 HAVSSC 9

RESULT 20
US-08-893-534A-32
Sequence 32, Application US/08893534A
Patent No. 6031072
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
TITLE OF INVENTION: CELL ADHESION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,534A
FILING DATE: 11-JUL-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /product= "OTHER"
OTHER INFORMATION: /note= "Residue is
US-08-893-534A-32

Query Match 78.6%; Score 33; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
Db 4 HAVSSC 9

RESULT 21
US-08-996-679-31
Sequence 31, Application US/08996679
Patent No. 6163071
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
TITLE OF INVENTION: CELL ADHESION
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,679
FILING DATE: 23-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 100086.401C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: circular
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /product= "OTHER"
OTHER INFORMATION: /note= "Residue is beta-mercaptopropionic acid"
US-08-996-679-31

Query Match 78.6%; Score 33; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
|||||
Db 4 HAVSSC 9

RESULT 22

US-08-996-679-32
; Sequence 32, Application US/08996679
; Patent No. 6169071

; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: CELL ADHESION
; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA

; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,679
; FILING DATE: 23-DEC-1997

CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 100086.401C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 32:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9' amino acid

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: circular

; MOLECULE TYPE: peptide

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 1

; OTHER INFORMATION: /product= "OTHER"

; OTHER INFORMATION: /note= "Residue is

; OTHER INFORMATION: beta,beta-pentamethylene-beta-mercaptopropionic acid"

US-08-996-679-32

Query Match 78.6%; Score 33; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
|||||
Db 4 HAVSSC 9

RESULT 23

US-09-115-395-47

; Sequence 47, Application US/09115395A

; Patent No. 6207639

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH

; FILE REFERENCE: 100086.401C3

; CURRENT APPLICATION NUMBER: US/09/115,395A

; CURRENT FILING DATE: 1998-07-14
; EARLIER APPLICATION NUMBER: 08/996,679
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/893,534
; EARLIER FILING DATE: 1997-07-11
; EARLIER APPLICATION NUMBER: 60/021,612
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Solid Phase

; OTHER INFORMATION: Synthesis

; FEATURE:

; NAME/KEY: MOD_RES

; LOCATION: (1)

; OTHER INFORMATION: Where Xaa is beta-mercaptopropionic acid

US-09-115-395-47

Query Match 78.6%; Score 33; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7

|||||

Db 4 HAVSSC 9

RESULT 24

US-09-115-395-48

; Sequence 48, Application US/09115395A

; Patent No. 6207639

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NEURITE OUTGROWTH

; FILE REFERENCE: 100086.401C3

; CURRENT APPLICATION NUMBER: US/09/115,395A

; CURRENT FILING DATE: 1998-07-14

; EARLIER APPLICATION NUMBER: 08/996,679

; EARLIER FILING DATE: 1997-12-23

; EARLIER APPLICATION NUMBER: 08/893,534

; EARLIER FILING DATE: 1997-07-11

; EARLIER APPLICATION NUMBER: 60/021,612

; EARLIER FILING DATE: 1996-07-12

; NUMBER OF SEQ ID NOS: 77

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 48

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Solid Phase

; OTHER INFORMATION: Synthesis

; FEATURE:

; NAME/KEY: MOD_RES

; LOCATION: (1)

; OTHER INFORMATION: Where Xaa is

; OTHER INFORMATION: beta,beta-pentamethylene-beta-mercaptopropionic

; OTHER INFORMATION: acid

US-09-115-395-48

Query Match 78.6%; Score 33; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7

|||||

Db 4 HAVSSC 9

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /product= "OTHER"
/note= "Residue is
beta,beta-pentamethylene-beta-mercaptopropionic acid"
SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-10-006-982-32

Query Match 78.6%; Score 33; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
Db 4 HAVSSC 9

RESULT 18

US-09-769-145-70
; Sequence 70, Application US/09769145
; Patent No. US20020168761A1
; GENERAL INFORMATION:
; APPLICANT: Gour, Barbara J.
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Ali, Ammar
; APPLICANT: Ni, Feng
; APPLICANT: Chen, Zhigang
; APPLICANT: Michaud, Stephanie
; APPLICANT: Wang, Shomeng
; APPLICANT: Hu, Zengjian
; TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
; FILE REFERENCE: 100086.413C1
; CURRENT APPLICATION NUMBER: US/09/769,145
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 09/491,078
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: Peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; NAME/KEY: MOD_RES
; LOCATION: (2)
; OTHER INFORMATION: Where Xaa is beta,beta-pentamethylene cysteine

Query Match 78.6%; Score 33; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
Db 5 HAVSSC 10

RESULT 19

US-10-058-821-49
; Sequence 49, Application US/10058821
; Publication No. US20030087811A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; APPLICANT: Fatoohi, Riaz
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
; FILE REFERENCE: 100086.401C12
; CURRENT APPLICATION NUMBER: US/10/058,821
; CURRENT FILING DATE: 2002-01-29
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
; OTHER INFORMATION: Peptide with classical cadherin cell adhesion
; OTHER INFORMATION: recognition sequence
; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
; OTHER INFORMATION: and/or C-terminal modifications such as amide or
; NAME/KEY: MOD_RES
; LOCATION: (2)
; OTHER INFORMATION: Where Xaa is beta,beta-pentamethylene cysteine

US-10-058-821-49

Query Match 78.6%; Score 33; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HAVSSC 7
Db 5 HAVSSC 10

RESULT 20

US-10-006-982-30
; Sequence 30, Application US/10006982
; Patent No. US20020151475A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
CELL ADHESION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
ADDRESS: SEED IP LAW GROUP PLLC
STREET: 6300 Bank of America Bldg., 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/006,982
FILING DATE: 04-Dec-2001
CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
NAME: Christiansen, William T.
REGISTRATION NUMBER: 44,614
REFERENCE/DOCKET NUMBER: 100086.401C11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: circular
MOLECULE TYPE: peptide
FEATURE:

NAME/KEY: Modified-site
LOCATION: 2
OTHER INFORMATION: /product= "OTHER"
/note= "Residue is beta,beta-pentamethylene cysteine"
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-10-006-982-30

Query Match 78.6%; Score 33; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.5; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0

QY 2 HAVSSC 7
DB 5 HAVSSC 10

RESULT 21

US-08-424-550B-59
Sequence 59, Application US/08424550B
Patent No. US20020119447A1
GENERAL INFORMATION:

APPLICANT: JOHN N. SIMONS
APPLICANT: TAMI J. PILOT-MATIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MURRHOF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-424-550B-59

Query Match 78.6%; Score 33; DB 8; Length 101;
Best Local Similarity 71.4%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 69 CHAVRXC 75

RESULT 22

US-10-047-542-72
Sequence 72, Application US/10047542
Patent No. US20020168367A1
GENERAL INFORMATION:

APPLICANT: LARRICK, JAMES W.
APPLICANT: WYCOFF, KEITH L.
TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL
TITLE OF INVENTION: AND BACTERIAL DISEASES
FILE REFERENCE: 030905.0004.CIP1
CURRENT APPLICATION NUMBER: US/10/047,542
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: PCT/US01/13932
PRIOR FILING DATE: 2001-04-28
PRIOR APPLICATION NUMBER: 60/200,298
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 72
LENGTH: 527
TYPE: PRT
ORGANISM: Cricetus griseus
US-10-047-542-72

Query Match 78.6%; Score 33; DB 9; Length 527;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVSS 6
DB 454 CHAVSS 459

RESULT 23

US-09-769-145-50
Sequence 50, Application US/09769145
Patent No. US20020168761A1
GENERAL INFORMATION:

APPLICANT: Gour, Barbara J.
APPLICANT: Blaschuk, Orest W.
APPLICANT: Ali, Ammar
APPLICANT: Ni, Peng
APPLICANT: Chen, Zhigang
APPLICANT: Michaud, Stephanie
APPLICANT: Wang, Shomeng
APPLICANT: Hu, Zengjian
TITLE OF INVENTION: PEPTIDOMIMETIC MODULATORS OF CELL ADHESION
FILE REFERENCE: 100086.413C1
CURRENT APPLICATION NUMBER: US/09/769,145
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 09/491,078
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 50
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Cyclic
OTHER INFORMATION: peptide with classical cadherin cell adhesion
OTHER INFORMATION: recognition sequence
OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group

OTHER INFORMATION: and/or C-terminal modifications such as amide or
 OTHER INFORMATION: ester group
 US-09-769-145-50

Query Match 76.2%; Score 32; DB 9; Length 7;
 Best Local Similarity 71.4%; Pred. No. 3.7e+05;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 Db 1 CHAVDIC 7

RESULT 24

US-10-058-821-31
 ; Sequence 31, Application US/10058821
 ; Publication No. US2003008781A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Blaschuk, Orest W.
 ; APPLICANT: Gour, Barbara J.
 ; APPLICANT: Farookhi, Riaz
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR CANCER THERAPY
 ; FILE REFERENCE: 10086.401C12
 ; CURRENT APPLICATION NUMBER: US/10/058,821
 ; CURRENT FILING DATE: 2002-01-29
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 31
 ; LENGTH: 7
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Cyclic
 ; OTHER INFORMATION: peptide with classical cadherin cell adhesion
 ; OTHER INFORMATION: recognition sequence
 ; OTHER INFORMATION: Cyclic Peptide may comprise N-terminal
 ; OTHER INFORMATION: modification such as acetyl or alkoxybenzyl group
 ; OTHER INFORMATION: and/or C-terminal modifications such as amide or
 ; OTHER INFORMATION: ester group
 US-10-058-821-31

Query Match 76.2%; Score 32; DB 9; Length 7;
 Best Local Similarity 71.4%; Pred. No. 3.7e+05;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 Db 1 CHAVDIC 7

RESULT 25

US-10-144-929-120
 ; Sequence 120, Application US/10144929
 ; Publication No. US20030069405A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: 70 Human Secreted Proteins
 ; FILE REFERENCE: P2014P1
 ; CURRENT APPLICATION NUMBER: US/10/144,929
 ; CURRENT FILING DATE: 2002-05-15
 ; PRIOR APPLICATION NUMBER: US/09/251,329
 ; PRIOR FILING DATE: 1999-02-17
 ; PRIOR APPLICATION NUMBER: PCT/US98/17044
 ; PRIOR FILING DATE: 1998-08-18
 ; NUMBER OF SEQ ID NOS: 257
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 120
 ; LENGTH: 39
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (39)

OTHER INFORMATION: Xaa equals stop translation
 US-10-144-929-120

Query Match 76.2%; Score 32; DB 9; Length 39;
 Best Local Similarity 57.1%; Pred. No. 42;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 Db 22 CHCMASC 28

Search completed: June 20, 2003, 21:02:49
 Job time : 23.9773 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:45:48 ; Search time 17.8182 seconds
(without alignments)
37.767 Million cell updates/sec

Title: US-10-105-008-44

Perfect score: 42

Sequence: 1 CHAVSSC 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : PIR 73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	88.1	320	2 A53119	cell adhesion glyco
2	37	88.1	643	2 S46723	arginine-trna liga
3	35	83.3	508	2 A83885	hypothetical prote
4	35	83.3	1785	2 A45546	major merizoite su
5	33	78.6	107	2 C91026	hypothetical prote
6	33	78.6	571	2 T33113	hypothetical prote
7	33	78.6	640	1 A30452	uromodulin precurs
8	33	78.6	642	1 S52111	uromodulin precurs
9	32	76.2	157	2 T17883	major capsid prote
10	32	76.2	214	2 S39831	hypothetical prote
11	32	76.2	277	2 A31840	RNA polymerase bet
12	32	76.2	393	1 A48357	nonstructural prot
13	32	76.2	1205	2 T18517	procollagen N-endo
14	31.5	75.0	702	2 D86469	protein F12K21.12
15	31	73.8	206	2 B89431	protein K02E2.6 [i
16	31	73.8	238	2 A43176	conserved hypothet
17	31	73.8	336	2 S55863	probable membrane
18	31	73.8	403	2 S07825	hypothetical prote
19	31	73.8	531	2 T51922	hypothetical prote
20	31	73.8	621	1 T01935	naringenin 3-dioxy
21	31	73.8	626	2 D88601	protein Y49E10.11
22	31	73.8	713	2 A35502	major surface-labe
23	31	73.8	1061	2 JC7116	Xsal-3 protein - A
24	31	73.8	1088	2 A69493	cysteine proteinas
25	31	73.8	1323	2 T30253	spalt protein - mo
26	31	73.8	1350	2 T30341	zinc finger protei
27	30	71.4	132	2 A84415	hypothetical prote
28	30	71.4	134	2 I48639	neurotoxin homolog
29	30	71.4	171	2 S09903	hypothetical prote

ALIGNMENTS

RESULT 1

A53119
cell adhesion glycoprotein gp64 precursor - slime mold (Polysphondylium pallidum)
C:Species: Polysphondylium pallidum
C>Date: 07-Jul-1995 #sequence revision 07-Jul-1995 #text_change 31-Jan-2000
C:Accession: A53119; S28836; P50434
R:Manabe, R.; Saito, T.; Kumazaki, T.; Sakaitani, T.; Nakata, N.; Ochiai, H.
J. Biol. Chem. 269, 528-535, 1994
A>Title: Molecular cloning and the COOH-terminal processing of Gp64, a putative cell
A:Reference number: A53119; MUID:94103266; PMID:8276846
A:Accession: A53119
A:Molecule type: mRNA; protein
A:Residues: 1-320 <MAN>
A:Cross-references: GB:D14993
R:Saito, T.; Kumazaki, T.; Ochiai, H.
Eur. J. Biochem. 211, 147-155, 1993
A>Title: A purification method and N-glycosylation sites of a 36-cysteine-containing
A:Reference number: S28836; MUID:93145941; PMID:8425525
A:Accession: S28836
A:Molecule type: protein
A:Residues: 20-299 <SAI>
R:Saito, T.; Kumazaki, T.; Ochiai, H.

hypothetical prote
protein C17B7.9 [i
hypothetical prote
hypothetical prote
NZ-3 antigen - hum
succinate dehydrog
succinate dehydrog
succinate dehydrog
succinate dehydrog
methionyl aminopep
methionine aminope
methionine aminope
hypothetical prote
glucuronosyltransf
transcription regu
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
UL53 protein - hum
DNA polymerase III
protein Y43F8A.2 [i
hypothetical prote
hypothetical prote
IMP dehydrogenase
hypothetical prote
hypothetical prote
VAC8 protein - yea
uromodulin precurs
Tamm-Horsfall prot
hypothetical prote
nuclear protein NO
hypothetical prote
adhesin Wf-1 - Aje
gastric mucin MUC5
adducin homolog -
MEGF6 protein - ra
resistance protein
resistance protein
hypothetical prote
cation-independent
probable guanine n
hypothetical prote
flavonol 3-sulfotr

submitted to JIPID, November 1992

A;Description: A purification method and N-glycosylation sites of a 18 cysteine containing
A;Reference number: PS0434
A;Accession: PS0434
A;Molecule type: protein
A;Residues: 20-41;48-53;78-107;134-144;151-162;181-192;207-219;285-299 <SA2>
A;Experimental source: strain WS320
C;Keywords: blocked carboxyl end; cell adhesion; duplication; glycoprotein; lipoprotein;
P;1-19/Domain: signal sequence #status predicted <SIG>
P;20-299/Product: cell surface glycoprotein gp64 #status experimental <MAT>
P;300-320/Domain: carboxyl-terminal propeptide #status predicted <CTP>
P;312,215,295/Binding site: carboxylate (Asn) (covalent) #status absent
P;49,80,141,158/Binding site: carboxylate (Asn) (covalent) #status experimental
P;187,216/Binding site: carboxylate (Asn) (covalent) #status atypical
P;299/Modified site: GSI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 88.1%; Score 37; DB 2; Length 320;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
Db 165 CHVSSC 171

RESULT 2

S46723
arginine-tRNA ligase (EC 6.1.1.19), mitochondrial - yeast (Saccharomyces cerevisiae)
N;Alternate names: arginyl-tRNA synthetase; protein H9332.9; protein YHR091C
C;Species: Saccharomyces cerevisiae
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 03-Jun-2002
C;Accession: S46723; S56044
R;Vaudin, M.

submitted to the EMBL Data Library, May 1994

A;Description: The sequence of S. cerevisiae cosmid 9332.

A;Reference number: S46715

A;Accession: S46723

A;Molecule type: DNA

A;Residues: 1-643 <VAU>

A;Cross-references: EMBL:U00060; NID:G487928; PIDN:AA68931.1; PID:G487937; MIPS:YHR091C

R;Tragoloff, A.A.; Shitanko, A.

submitted to the EMBL Data Library, January 1995

A;Description: Yeast MSRI gene.

A;Reference number: S56044

A;Accession: S56044

A;Molecule type: DNA

A;Residues: 1-498, T, 500-643 <TZA>

A;Cross-references: EMBL:L39019; NID:G633021; PIDN:AAA61486.1; PID:G633022

C;Genetics:

A;Gene: SGD:MSR1

A;Cross-references: SGD:S0001133; MIPS:YHR091C

A;Map position: 8R

A;Genome: nuclear

C;Superfamily: arginine-tRNA ligase

C;Keywords: aminoacyl-tRNA synthetase; ligase; mitochondrion; protein biosynthesis

Query Match 88.1%; Score 37; DB 2; Length 643;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
Db 594 CHQVSSC 600

RESULT 3

A83885

hypothetical protein BH1881 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: A83885

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AP001513; GB:BA000004; NID:G10174345; PIDN:BA05600.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1881

Query Match 83.3%; Score 35; DB 2; Length 508;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
Db 427 CHAVGAC 433

RESULT 4

A45546

major merozoite surface antigen precursor - Plasmodium chabaudi chabaudi
C;Species: Plasmodium chabaudi chabaudi
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C;Accession: A45546

R;Deleersnijder, W.; Hendrix, D.; Bendahman, N.; Haneegreets, J.; Brijs, L.; Hamers-Ca
MOl. Biochem. Parasitol. 43, 231-244, 1990

A;Title: Molecular cloning and sequence analysis of the gene encoding the major merozo
A;Reference number: A45546; MUID:91218805; PMID:2090945

A;Accession: A45546

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1785

A;Cross-references: GB:M34947; NID:G160597; PID:G160598

C;Superfamily: major merozoite surface antigen

C;Keywords: surface antigen

Query Match 83.3%; Score 35; DB 2; Length 1785;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||
Db 422 CHSTSSC 428

RESULT 5

C91026

hypothetical protein ECs3179 [imported] - Escherichia coli (strain O157:H7, substrain

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C;Accession: C91026

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gaasawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C91026

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-107 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA836602.1; PID:G13362649; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs3179

Query Match 78.6%; Score 33; DB 2; Length 107;
Best Local Similarity 57.1%; Pred. No. 44;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
|||

Db 95 CHAITPC 101

RESULT 6

T33113

hypothetical protein B0511.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Nov-2000

C:Accession: T33113

R:Tin-Wollam, A.; Sutterer, C.; Ozersky, P.

submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid B0511.

A:Reference number: Z21285

A:Accession: T33113

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-571 <TIN>

A:Cross-references: EMBL:AF067608; PIDN:AAAC17654.1; GSPDB:GN00019; CESP:B0511.6

A:Experimental source: strain Bristol N2; clone B0511

C:Genetics:

A:Gene: CESP:B0511.6

A:Map position: 1
A:Introns: 104/3; 345/3; 446/3; 485/3; 512/3; 546/3

Query Match 78.6%; Score 33; DB 2; Length 571;

Best Local Similarity 71.4%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

|||::|

Db 499 CHGVSR 505

RESULT 7

A30452

uromodulin precursor - human

N:Alternate names: Tamm-Horsfall urinary glycoprotein; uromucoid

C:Species: Homo sapiens (man)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A30452; A30453

R:Pennica, D.; Kohr, W.J.; Kuang, W.J.; Glaister, D.; Aggarwal, B.B.; Chen, E.Y.; Goedde

Science 236, 83-88, 1987

A:Title: Identification of human uromodulin as the Tamm-Horsfall urinary glycoprotein.

A:Reference number: A30452; MUID:87177970; PMID:3453112

A:Accession: A30452

A:Molecule type: mRNA

A:Residues: 1-640 <PEN>

A:Cross-references: GB:M15881; NID:q340163; PIDN:AAA36798.1; PID:q340164

A:Note: parts of this sequence, including the amino end of the mature protein, were dete

R:Heslon, C.; Decker, J.M.; Sherblom, A.P.; Kumar, S.; Yue, C.C.; Mattaliano, R.J.; Tiza

Science 237, 1479-1484, 1987

A:Title: Uromodulin (Tamm-Horsfall glycoprotein): A renal ligand for lymphokines.

A:Reference number: A30453; MUID:87319675; PMID:3498215

A:Accession: A30453

A:Molecule type: mRNA

A:Residues: 1-640 <HES>

A:Cross-references: GB:M1778

A:Note: the authors translated the codon GTC for residue 381 as Asp

A:Note: parts of this sequence, including the amino end of the mature protein, were dete

R:Rindler, M.J.; Naik, S.S.; Li, N.; Hoops, T.C.; Peraldi, M.N.

J. Biol. Chem. 265, 20784-20789, 1990

A:Title: Uromodulin (Tamm-Horsfall glycoprotein/uromucoid) is a phosphatidylinositol-lin

A:Reference number: A30454; MUID:91065873; PMID:2249987

A:Contents: annotation; GPI-anchor

C:Comment: The protein is anchored to the cell membrane by a phosphatidylinositol linkag

e membrane-bound form.

C:Genetics:

A:Gene: GDB:UMOD

A:Cross-references: GDB:133729; OMIM:191845

A:Map position: 16p13.11-16p12.3

A:Introns: 30/1; 289/1; 325/1; 394/3; 444/2; 526/2; 580/3; 608/1; 621/1

A:Note: the first intron occurs before the initiator codon.

C:Superfamily: uromodulin; EGF homology; membrane glycoprotein 2 homology; ZP domain hom

C:Keywords: duplication; glycoprotein; lipoprotein; phosphatidylinositol linkage; ph
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-640/Product: uromodulin #status predicted <MAT>
F:69-106/Domain: EGF homology <EGF>

F:170-639/Domain: membrane glycoprotein 2 homology <MGH>

F:334-583/Domain: ZP domain homology <ZPH>

F:625-640/Domain: transmembrane #status predicted <TRM>

F:38,76,80,275,322,513/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:232,356/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 78.6%; Score 33; DB 1; Length 640;

Best Local Similarity 57.1%; Pred. No. 1.7e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

|||::|

Db 120 CHALATC 126

RESULT 8

S52111

uromodulin precursor - mouse

N:Alternate names: Tamm-Horsfall urinary glycoprotein; uromucoid

C:Species: Mus musculus (house mouse)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S52111

R:Prasadnan, K.; Bates, J.; Badgett, A.; Dell, M.; Sukhatme, V.; Yu, H.; Kumar, S.

Biochim. Biophys. Acta 1260, 328-332, 1995

A:Title: Nucleotide sequence and peptide motifs of mouse uromodulin (Tamm-Horsfall p

A:Reference number: S52111; MUID:95178555; PMID:7873609

A:Accession: S52111

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-642 <PRA>

C:Superfamily: uromodulin; EGF homology; membrane glycoprotein 2 homology; ZP domain

C:Keywords: duplication; glycoprotein; lipoprotein; membrane protein; phosphatidylin

F:69-105/Domain: EGF homology <EGL>

F:111-147/Domain: EGF homology <EG2>

F:171-641/Domain: membrane glycoprotein 2 homology <MGH>

F:335-584/Domain: ZP domain homology <ZPH>

Query Match 78.6%; Score 33; DB 1; Length 642;

Best Local Similarity 57.1%; Pred. No. 1.7e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

|||::|

Db 119 CHALATC 125

RESULT 9

T17883

major capsid protein homolog A382R - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17883

R:Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z18806

A:Accession: T17883

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-157 <GRA>

A:Cross-references: EMBL:U42580; NID:q4028896; PIDN:AAC96750.1

A:Experimental source: specific host Chlorella strain NC64A

C:Genetics:

A:Note: A382R

Query Match 76.2%; Score 32; DB 2; Length 157;

Best Local Similarity 71.4%; Pred. No. 89;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

Db 102 CHAVRRC 108
|||||

RESULT 10

S39831
hypothetical protein YBL057c - yeast (Saccharomycetes cerevisiae)
N;Alternate names: hypothetical protein YBL0510; hypothetical protein YBL0514
C;Species: Saccharomycetes cerevisiae
C;Date: 16-Dec-1993 #sequence_revision 09-Sep-1994 #text_change 19-Apr-2002
C;Accession: S39831; S45792; S37332
R;Scherens, B.; el Bakkoury, M.; Vierendeels, F.; Dubois, E.; Messenguy, F.
Yeast 9, 1355-1371, 1993
A;Title: Sequencing and functional analysis of a 32 560 bp segment on the left arm of ye
A;Reference number: S39824; MUID:94205266; PMID:8154187
A;Accession: S39831
A;Molecule type: DNA
A;Residues: 1-214 <SCH>
A;Cross-references: EMBL:Z23261; NID:g313733; PIDN:CAA80790.1; PID:g313741
A;Experimental source: strain S288C
R;Dubois, E.; El Bakkoury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens, B.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45782
A;Accession: S45792
A;Molecule type: DNA
A;Residues: 1-214 <DUB>
A;Cross-references: EMBL:Z35818; NID:g536087; PIDN:CAA84877.1; PID:g536088; MIPS:YBL057c
A;Experimental source: strain S288C
C;Genetics:
A;Cross-references: SGD:S0000153
A;Map position: 2L

Query Match 76.2%; Score 32; DB 2; Length 214;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
||| |||
Db 116 CHAALSC 122

RESULT 11

A31840
RNA polymerase beta chain homolog - Leptospira biflexa (serotype patoc)
N;Alternate names: asge-complementing protein
C;Species: Leptospira biflexa
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 08-Oct-1999
C;Accession: A31840
R;Zuerner, R.L.; Charon, N.W.
J. Bacteriol. 170, 4548-4554, 1988
A;Title: Nucleotide sequence analysis of a gene cloned from Leptospira biflexa serovar P
A;Reference number: A31840; MUID:89008067; PMID:2844724
A;Accession: A31840
A;Molecule type: DNA
A;Residues: 1-277 <ZUE>
A;Cross-references: GB:M22622; NID:g149608; PIDN:AAA25260.1; PID:g149609

Query Match 76.2%; Score 32; DB 2; Length 277;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
||| |||
Db 265 CHTLSTC 271

RESULT 12

A48357
nonstructural protein NS53 - porcine rotavirus C (strain Cowden)
C;Species: porcine rotavirus C
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C;Accession: A48357; S16051
R;Bremont, M.; Chabanne-Vautherot, D.; Cohen, J.

Arch. Virol. 130, 85-92, 1993
A;Title: Sequence analysis of three non structural proteins of a porcine group C (Cow
A;Reference number: A48357; MUID:93277387; PMID:8389118
A;Accession: A48357
A;Molecule type: genomic RNA
A;Residues: 1-393 <BRE>
A;Cross-references: GB:X60546; NID:g61348; PIDN:CAA43036.1; PID:g61349
A;Experimental source: strain Cowden
A;Note: sequence extracted from NCBI backbone (NCBIN:132688, NCBIIP:132689)
R;Bremont, M.; Brottier, P.; Chabanne-Vautherot, D.; Cohen, J.
submitted to the EMBL Data Library, July 1991
A;Description: "zinc finger" motif conservation in a group C Rotavirus genome.
A;Reference number: S16051
A;Accession: S16051
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-393 <BRM>
A;Cross-references: EMBL:X60546; NID:g61348; PIDN:CAA43036.1; PID:g61349
C;Genetics:
A;Map position: segment 7
C;Superfamily: porcine rotavirus C nonstructural protein NS53
C;Keywords: glycoprotein; nonstructural protein; zinc finger
F;39-68/Region: zinc finger
F;290/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 76.2%; Score 32; DB 1; Length 393;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
||| |||
Db 273 CHTLSC 279

RESULT 13

T18517
procollagen N-endopeptidase (EC 3.4.24.14) I - bovine
N;Alternate names: procollagen N-proteinase
C;Species: Bos primigenius taurus (cattle)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18517
R;Collige, A.; Nusgens, B.V.; Lapiere, C.M.
submitted to the EMBL Data Library, February 1996
A;Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.
A;Reference number: Z18941
A;Accession: T18517
A;Status: preliminary; translated from GB/EMBL/DDRU
A;Molecule type: mRNA
A;Residues: 1-1205 <COL>
A;Cross-references: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CAA65253.1
A;Experimental source: skin
C;Genetics:
A;Gene: PC I-NP
C;Function:
A;Description: catalyzes cleavage of the propeptides of type I and II collagens prior
C;Keywords: hydrolase; metalloproteinase

Query Match 76.2%; Score 32; DB 2; Length 1205;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
||| |||
Db 380 CHVRSC 386

RESULT 14

D86469
protein P12K21.12 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D86469
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Aloni

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Majti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A6141; MUID:21016719; PMID:11130712
A:Accession: D86469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-702 <STO>
A:Cross-references: GB:AE005172; NID:g8778248; PIDN:AAF79257.1; GSPDB:GN00141
C:Genetics:
A:Gene: F12K21.12
A:Map position: 1

Query Match 75.0%; Score 31.5; DB 2; Length 702;
Best Local Similarity 87.5%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CH-AVSSC 7
|||
437 CHIAVSSC 444

Db

RESULT 15
B89431
protein K02E2.6 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B89431
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B89431
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <STO>
A:Cross-references: GB:chr_V; PIDN:CAB04550.1; PID:g3878199; GSPDB:GN00023; CESF:K02E2.6
C:Genetics:
A:Gene: K02E2.6
A:Map position: 5

Query Match 73.8%; Score 31; DB 2; Length 206;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
|||
132 CHPISHC 138

Db

RESULT 16
AH3176
conserved hypothetical protein Atu5140 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AH3176
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AH3176

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL45830.1; PID:g17743570; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu5140
A:Genome: plasmid

Query Match 73.8%; Score 31; DB 2; Length 238;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
|||
31 CHELSAC 37

Db

RESULT 17
S55863
probable membrane protein YNL326c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N0325
C:Species: Saccharomyces cerevisiae
C:Date: 27-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 19-Apr-2002
C:Accession: S55863; S51289; S63307
R:Matthi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
Yeast 11, 567-572, 1995
A:Title: Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV identifie
A:Reference number: S55859; MUID:95373280; PMID:7645347
A:Accession: S55863
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-336 <MAF>
A:Cross-references: EMBL:Z46259; NID:G633655; PIDN:CAA86372.1; PID:G633660
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R:Nicaud, J.J.
submitted to the EMBL Data Library, January 1995
A:Description: Sequence analysis of a 13.9 Kb fragment of yeast chromosome XIV ident
A:Reference number: S51285
A:Accession: S51289
A:Molecule type: DNA
A:Residues: 1-336 <NIC>
A:Cross-references: EMBL:Z46259; NID:G633655; PID:G633660
R:Matthi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63287
A:Accession: S63307
A:Molecule type: DNA
A:Residues: 1-336 <MAW>
A:Cross-references: EMBL:Z71602; NID:G1302442; PID:g239770; PID:g1302443; MIPS:YNL32
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0005270
A:Map position: 14L
C:Keywords: transmembrane protein
F:37-53/Domain: transmembrane #status predicted <TM1>
F:149-165/Domain: transmembrane #status predicted <TM2>
F:189-205/Domain: transmembrane #status predicted <TM3>

Query Match 73.8%; Score 31; DB 2; Length 336;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
|||
117 CHHCSSC 123

Db

RESULT 18
S07825
hypothetical protein 2 - fruit fly (Drosophila melanogaster) transposon FB
C:Species: Drosophila melanogaster
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 21-Jul-2000

C;Accession: S07825
R;Templeton, N.S.; Potter, S.S.
EMBO J. 8, 1887-1894, 1989
A;Title: Complete foldback transposable elements encode a novel protein found in Drosophila
A;Reference number: S07824; MUID:89356666; PMID:2548860
A;Accession: S07825
A;Molecule type: DNA
A;Residues: 1-403 <TEM>
A;Cross-references: EMBL:X15469; NID:g7962; PIDN:CAA33497.1; PID:g7964
C;Genetics:
A;Gene: FlyBase:NOF
A;Cross-references: FlyBase:FBgn0002949
C;Keywords: nucleus

Query Match 73.8%; Score 31; DB 2; Length 403;
Best Local Similarity 57.1%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CHAVSSC 7
Db 336 CHRICSC 342

RESULT 19

T51922
hypothetical protein B23111.340 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T51922
R;Schulte, U.; Aign, V.; Hohnel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura, R.
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25958
A;Accession: T51922
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531 <SCH>
A;Cross-references: EMBL:AL391572; GSPDB:GN00116; NCSP:B23111.340
A;Experimental source: BAC clone B23111; strain OR74A
C;Genetics:
A;Gene: NCSP:B23111.340
A;Map position: 6
A;Introns: 42/3; 439/3; 505/2

Query Match 73.8%; Score 31; DB 2; Length 531;
Best Local Similarity 57.1%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CHAVSSC 7
Db 6 CHSTESC 12

RESULT 20

T01935
naringenin 3-dioxygenase (EC 1.14.11.9) - common tobacco
N;Alternate names: Flavonone 3-hydroxylase
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-May-2000
C;Accession: T01935
R;Kim, G.; Kim, Y.
Plant Physiol. 116, 1605, 1998
A;Title: Flavonone 3-hydroxylase (Accession No. AF036093) cloning and sequencing from N. glauca
A;Reference number: Z14463
A;Accession: T01935
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-621 <KIM>
A;Cross-references: EMBL:AF036093; NID:g2828005; PIDN:AAC15414.1; PID:g2828006
A;Experimental source: tissue-type petal
C;Superfamily: tobacco naringenin 3-dioxygenase
C;Keywords: ascorbic acid; oxidoreductase

Query Match 73.8%; Score 31; DB 1; Length 621;

Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CHAVSSC 7
Db 366 CHPVSV 372

RESULT 21
D88601
protein Y49E10.11 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: D88601
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A;Accession: D88601
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-626 <STO>
A;Cross-references: GB:chr_III; PIDN:CAB11550.1; PID:g3979990; GSPDB:GN00021; CESP:Y4
C;Genetics:
A;Gene: Y49E10.11
A;Map position: 3

Query Match 73.8%; Score 31; DB 2; Length 626;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 94 CHAVVCC 100

RESULT 22

A35502
major surface-labeled trophozoite antigen precursor - Giardia lamblia
C;Species: Giardia lamblia
C;Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 09-Sep-1997
C;Accession: A35502
R;Gill, F.D.; Hagblom, P.; Harwood, J.; Aley, S.B.; Reiner, D.S.; McCaffery, M.; So
Proc. Natl. Acad. Sci. U.S.A. 87, 4463-4467, 1990
A;Title: Isolation and expression of the gene for a major surface protein of Giardia
A;Reference number: A35502; MUID:90280395; PMID:2352929
A;Accession: A35502
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-713 <GIL>
A;Cross-references: GB:M33641; NID:gl59131; PID:gl59132
C;Keywords: surface antigen; transmembrane protein

Query Match 73.8%; Score 31; DB 2; Length 713;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 331 CEAUSNC 337

RESULT 23

JC7116
Xsal-3 protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C;Accession: JC7116
R;Onuma, Y.; Nishinakamura, R.; Takahashi, S.; Yokota, T.; Asashima, M.
Biochem. Biophys. Res. Commun. 264, 151-156, 1999
A;Title: Molecular cloning of a novel Xenopus spalt gene (Xsal-3).

A;Reference number: JC7116; MUID:99458636; PMID:10527856

A;Accession: JC7116

A;Molecule type: DNA

A;Residues: 1-1061 <ONU>

A;Cross-references: DDBJ:AB030827

A;Experimental source: egg, neural tube

C;Genetics:

A;Gene: Xsal-3

C;Keywords: egg; zinc finger

Query Match 73.8%; Score 31; DB 2; Length 1061;

Best Local Similarity 71.4%; Pred. No. 5.6e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7

Db 542 CHRVLSC 548

RESULT 24

A69493

cysteine proteinase homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C;Accession: A69493

R;Klenk, H.P.; Claytor, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: A69493

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1088 <KLE>

A;Cross-references: GB:AE000969; GB:AE000782; NID:92689292; PIDN:AA889309.1; PID:g264859

Query Match

Best Local Similarity 71.4%; Score 31; DB 2; Length 1088;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7

Db 132 CHDFSC 138

RESULT 25

T30253

spalt protein - mouse (fragment)

N;Alternate names: zinc finger protein meal

C;Species: Mus musculus (house mouse)

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000

C;Accession: T30253

R;Ott, T.; Kaetner, K.H.; Monaghan, A.P.; Schutz, G.

Mech. Dev. 56, 117-128, 1996

A;Title: The mouse homolog of the region specific homeotic gene spalt of Drosophila is e

A;Reference number: Z20791; MUID:96391179; PMID:8798152

A;Accession: T30253

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1323 <OTT>

A;Cross-references: EMBL:X97581; NID:gl296844; PIDN:CAA66196.1; PID:gl296845

C;Genetics:

A;Gene: msal

C;Function:

A;Description: may play an important role in the development of the nervous system

Query Match

Best Local Similarity 71.4%; Score 31; DB 2; Length 1323;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7

Db 668 CHRVLSC 674

Search completed: June 20, 2003, 20:59:18

Job time: 20.8182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:33:08 ; Search time 8.43182 Seconds
(without alignments)
34.433 Million cell updates/sec

Title: US-10-105-008-44
Perfect score: 42
Sequence: 1 CHAVSSC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	88.1	643	1 SYRM_YEAST	P38714 saccharomyc
2	33	78.6	587	1 GOX_PENAG	P81156 penicillium
3	33	78.6	605	1 GOX_TALFL	Q92452 talaromyc
4	33	78.6	640	1 UROM_HUMAN	P07911 homo sapien
5	33	78.6	643	1 UROM_BOVIN	P48733 bos taurus
6	32	76.2	214	1 YBF7_YEAST	P34222 saccharomyc
7	32	76.2	277	1 ARG8_LEPBI	P13440 leptospira
8	32	76.2	393	1 VN53_ROTFC	Q00033 porcine rot
9	32	76.2	1205	1 ATS2_BOVIN	P79331 b adams-2
10	32	76.2	1205	1 ATS3_HUMAN	O15072 homo sapien
11	32	76.2	1211	1 ATS2_HUMAN	O95450 h adams-2
12	31	73.8	88	1 YOR8_TTV1	P19283 thermoprote
13	31	73.8	336	1 YN66_YEAST	P42836 saccharomyc
14	31	73.8	403	1 NQF2_DROME	P15297 drosophila
15	31	73.8	713	1 TSA4_GIALA	P21849 giardia lam
16	31	73.8	1300	1 SAL3_HUMAN	Q9bxa9 homo sapien
17	31	73.8	1322	1 SAL1_MOUSE	Q9er74 mus musculu
18	31	73.8	1323	1 SAL3_MOUSE	Q62255 mus musculu
19	31	73.8	1324	1 SAL1_HUMAN	Q9nsc2 homo sapien
20	30	71.4	134	1 LY6F_MOUSE	P35460 mus musculu
21	30	71.4	171	1 IR10_HCMVA	P16808 human cytom
22	30	71.4	238	1 DHSB_ECOLI	P07014 escherichia
23	30	71.4	238	1 DHSB_SALTY	Q8zq2 salmonella
24	30	71.4	264	1 AMPM_ECOLI	P07906 escherichia
25	30	71.4	376	1 UL53_HCMVA	P16794 human cytom
26	30	71.4	534	1 UD15_HUMAN	P35504 homo sapien
27	30	71.4	578	1 VAC8_YEAST	P39968 saccharomyc
28	30	71.4	589	1 ENCL_HUMAN	O14682 homo sapien
29	30	71.4	589	1 ENCL_MOUSE	O35709 mus musculu
30	30	71.4	644	1 UROM_RAT	P27590 rattus norv-
31	30	71.4	685	1 NOP4_YEAST	P37838 saccharomyc
32	30	71.4	812	1 AXN2_BRARE	P57095 brachydanio
33	30	71.4	841	1 PHL2_HUMAN	Q15127 homo sapien

34	30	71.4	1156	1 HTS_DROME	Q02645 drosophila
35	29	69.0	140	1 DHSB_ARATH	P21915 arabidopsis
36	29	69.0	143	1 DHSB_RAT	P21913 rattus norv
37	29	69.0	180	1 Y335_TREPA	O83355 treponema p
38	29	69.0	239	1 DHSB_FORPU	P80477 porphyra pu
39	29	69.0	245	1 KSU5_ECOLI	P42216 escherichia
40	29	69.0	250	1 DHSB_CHOCR	P48932 chondrus cr
41	29	69.0	252	1 DHSB_SCHPO	P21911 schizosacch
42	29	69.0	259	1 DHSB_PARDE	Q59662 paracoccus
43	29	69.0	261	1 DHSB_RICGN	Q92178 rickettsia
44	29	69.0	261	1 DHSB_RICPR	Q92178 rickettsia
45	29	69.0	262	1 DHSB_CYACA	P48933 cyanidium c
46	29	69.0	263	1 AMPM_SALTY	P10882 salmonella
47	29	69.0	266	1 DHSB_YEAST	P21801 saccharomyc
48	29	69.0	280	1 DHSB_HUMAN	P21912 homo sapien
49	29	69.0	281	1 YH71_AERPE	Q9Y024 aeropyrum p
50	29	69.0	295	1 DHSB_DSTWA	P32420 ustilago ma
51	29	69.0	297	1 DHSB_DROME	P21914 drosophila
52	29	69.0	298	1 DHSB_CABEL	Q09545 caenorhabdi
53	29	69.0	311	1 F3ST_FLACH	P52836 flaveria ch
54	29	69.0	312	1 F3ST_PLABI	P52835 flaveria bi
55	29	69.0	323	1 VALI_PASVK	Q00338 panicum str
56	29	69.0	346	1 RECA_THIFE	P16238 thiobacillu
57	29	69.0	351	1 OPSP_CHICK	P51475 gallus gall
58	29	69.0	363	1 ASK1_ARATH	P43291 arabidopsis
59	29	69.0	413	1 Y4TS_RHISN	P55519 rhizobium s
60	29	69.0	415	1 DCDA_HABIN	P44316 haemophilus
61	29	69.0	416	1 GLYA_HELPF	Q9amp7 helicobacte
62	29	69.0	416	1 GLYA_HELPF	P56089 helicobacte
63	29	69.0	432	1 AGAL_BACSU	O34645 bacillus su
64	29	69.0	478	1 AMY_STRHY	P08486 streptomyce
65	29	69.0	481	1 YRDD_CABEL	Q18953 caenorhabdi
66	29	69.0	528	1 ORDA_ASPFL	P79084 aspergillus
67	29	69.0	543	1 VIBE_VIBCH	O07899 vibrio chol
68	29	69.0	611	1 LEM2_CANFA	P33730 canis fami
69	29	69.0	630	1 ESAB_TRYBB	P23799 trypanosoma
70	29	69.0	630	1 ESAB_TRYEQ	P26337 trypanosoma
71	29	69.0	706	1 HDAL_YEAST	P53973 saccharomyc
72	29	69.0	851	1 OBP_HSV11	P10193 herpes simp
73	29	69.0	955	1 TSP4_XENLA	Q06441 xenopus lae
74	29	69.0	982	1 YS96_CABEL	Q09965 caenorhabdi
75	29	69.0	1133	1 RAD8_SCHPO	P36607 schizosacch

ALIGNMENTS

RESULT 1
ID SYRM_YEAST STANDARD; PRT; 643 AA.
AC P38714;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Arginyl-tRNA synthetase, mitochondrial precursor (EC 6.1.1.19)
DE (Arginine-tRNA ligase) (ArgRS).
GN MSRI OR YHR091C.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Tzagoloff A.A.; Shtanko A.;
RL Submitted (XX-1995) to the EMBL/GenBank/DBJ databases.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,

RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaubin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII.";
 RL Science 265:2077-2082(1994).
 CC -!- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +
 CC diphosphate + L-arginyl-tRNA(Arg).
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L39019; AAA61486.1; -;
 DR EMBL; U00060; AAB68931.1; -;
 DR PIR; S46723; S46723.
 DR HSP; Q05506; IBS2.
 DR SGD; S0001133; MSRI.
 DR InterPro; IPR001278; Arg tRNA-synt_1c.
 DR InterPro; IPR005148; N_
 DR InterPro; IPR001412; tRNA-synt I.
 DR Pfam; PF00750; tRNA-synt_1d; 1.
 DR Pfam; PF03485; N-Arg; 1.
 DR PRINTS; PRO1038; TRNASYNTHARG.
 DR TIGRFAMs; TIGR00456; args; 1.
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Mitochondrion; Transit peptide.
 FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 643 ARGINYL-TRNA SYNTHETASE.
 FT SITE 188 198 "HIGH" REGION.
 FT CONFLICT 499 499 T -> M (IN REF. 2).
 SQ SEQUENCE 643 AA; 73663 MW; EFB4B5D25A18CB9 CRC64;
 Query Match 88.1%; Score 37; DB 1; Length 643;
 Best Local Similarity 85.7%; Pred. No. 9.6;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 DB 594 CHQVSSC 600
 RESULT 2
 GOX PENAG STANDARD; PRT; 587 AA.
 ID GOX PENAG STANDARD; PRT; 587 AA.
 AC P81156;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Glucose oxidase (EC 1.1.3.4) (Glucose oxyhydrase) (GOD)
 DE (Beta-D-glucose: oxygen 1-oxido-reductase).
 OS Penicillium amagasakiense.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
 RX NCBI_TaxID=63559;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=98181890; PubMed=9523716;
 RA Kiess M., Hecht H.-J., Kalisz H.M.;
 RT "Glucose oxidase from Penicillium amagasakiense. Primary structure
 RT and comparison with other glucose-methanol-choline (GMC)
 RT oxidoreductases.";
 RL Eur. J. Biochem. 252:90-99(1998).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).

RX MEDLINE=99234348; PubMed=10216293;
 RA Wohlfahrt G., Witt S., Hendle J., Schomburg D., Kalisz H.M.,
 RA Hecht H.-J.;
 RT "1.8 and 1.9-A resolution structures of the Penicillium amagasakiense
 RT and Aspergillus niger glucose oxidases as a basis for modelling
 RT substrate complexes";
 RL Acta Crystallogr. D 55:969-977(1999).
 CC -!- CATALYTIC ACTIVITY: Beta-D-glucose + O(2) = D-glucono-1,5-lactone
 CC + H(2O)(2).
 CC -!- COFACTOR: FAD.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THIS ENZYME IS WIDELY APPLIED FOR THE DETERMINATION
 CC OF GLUCOSE IN BODY FLUIDS AND IN REMOVING RESIDUAL GLUCOSE OR
 CC OXYGEN FROM FOODS AND BEVERAGES. FURTHERMORE, GLUCOSE
 CC OXIDASE-PRODUCING MOULDS SUCH AS ASPERGILLUS AND PENICILLIUM
 CC SPECIES ARE USED FOR THE BIOLOGICAL PRODUCTION OF GLUCONIC ACID.
 CC -!- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.
 DR PDB; 1GPE; 06-MAY-99.
 DR InterPro; IPR000172; GMC oxred.
 DR Pfam; PF00732; GMC oxred; 1.
 DR PROSITE; PS00623; GMC OXRED 1; 1.
 DR PROSITE; PS00624; GMC OXRED 2; 1.
 KW Oxidoreductase; Flavoprotein; FAD; Glycoprotein; 3D-structure.
 FT NP_BIND 26 55 POTENTIAL.
 FT ACT_SITE 525 525
 FT DISULFID 168 210
 FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 SQ SEQUENCE 587 AA; 63965 MW; 2827477B3DF4508C CRC64;
 Query Match 78.6%; Score 33; DB 1; Length 587;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HAVSSC 7
 DB 520 HAVSSC 525
 RESULT 3
 GOX TALPL STANDARD; PRT; 605 AA.
 ID GOX TALPL STANDARD; PRT; 605 AA.
 AC Q92452;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Glucose oxidase precursor (EC 1.1.3.4) (Glucose oxyhydrase) (GOD)
 DE (Beta-D-glucose: oxygen 1-oxido-reductase).
 GN GOX.
 OS Talaromyces flavus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Talaromyces.
 OX NCBI_TaxID=5095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=ATCC 32908;
 RX MEDLINE=98041929; PubMed=9371889;
 RA Murray F.R., Llewellyn D.J., Peacock W.J., Dennis E.S.;
 RT "Isolation of the glucose oxidase gene from Talaromyces flavus and
 RT characterisation of its role in the biocontrol of Verticillium
 RT dahliae.";
 RL Curr. Genet. 32:367-375(1997).
 CC -!- CATALYTIC ACTIVITY: Beta-D-glucose + O(2) = D-glucono-1,5-lactone
 CC + H(2O)(2).
 CC -!- COFACTOR: FAD.
 CC -!- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

or send an email to license@iab-sib.ch.

EMBL; U56240; AAB09442.1; --
HSP; P81156; 1GPE
InterPro; IPR000172; GMC_oxred.
Pfam; PF00732; GMC_oxred; 1.
PROSITE; PS00623; GMC_OXRED 1; 1.
PROSITE; PS00624; GMC_OXRED 2; 1.
Oxidoreductase; Flavoprotein; PAD; Glycoprotein; Signal.
SIGNAL 1 22 POTENTIAL:
CHAIN 23 605
NP_BIND 44 73
ACT_SITE 543 543
FAD (ADP PART) (PROBABLE).
POTENTIAL:
FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 375 375 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 519 519 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 605 AA; 65769 MW; 746F5F77FD1558D3 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7

Db 538 HAVSSC 543

RESULT 4

UROM_HUMAN
ID UROM_HUMAN STANDARD; PRT; 640 AA.
AC P07911;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Uromodulin precursor (Tamm-Horsfall urinary glycoprotein) (THP).
GN UMOD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87177970; PubMed=3453112;
RA Pennica D., Kohr W.J., Kuang W.-J., Glaister D., Aggarwal B.B.,
Chen E.Y., Goeddel D.V.;
RT "Identification of human uromodulin as the Tamm-Horsfall urinary glycoprotein."
RL Science 236:83-88(1987).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=87319675; PubMed=3498215;
RA Hession C., Decker J.M., Sherblom A.P., Kumar S., Yue C.C.,
Mataliano R.J., Tizard R., Kawashima E., Schmeissner U.,
Halecky S., Chow E.P., Burne C.A., Shaw A., Muchmore A.V.;
RT "Uromodulin (Tamm-Horsfall glycoprotein): a renal ligand for lymphokines."
RL Science 237:1479-1484(1987).
RN [3]
RP GPI-ANCHOR.
RX MEDLINE=91065873; PubMed=2249987;
RA Rindler M.J., Naik S.S., Li N., Hoops T.C., Peraldi M.-N.;
RT "Uromodulin (Tamm-Horsfall glycoprotein/uromucoicid) is a phosphatidylinositol-linked membrane protein."
RL J. Biol. Chem. 265:20784-20789(1990).
CC -1- FUNCTION: NOT KNOWN. MAY PLAY A ROLE IN REGULATING THE CIRCULATING ACTIVITY OF CYTOKINES AS IT BINDS TO IL-1, IL-2 AND TNF WITH HIGH AFFINITY.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR, THEN CLEAVED TO PRODUCE A SOLUBLE FORM WHICH IS SECRETED IN

CC -1- TISSUE SPECIFICITY: SYNTHESIZED BY THE KIDNEYS AND IS THE MOST ABUNDANT PROTEIN IN NORMAL HUMAN URINE.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.iab-sib.ch/announce/> or send an email to license@iab-sib.ch).

EMBL; M15881; AAA36798.1; --
EMBL; M17778; AAA36799.1; --
PIR; A30452; A30452.
HSP; P07204; IADX.
GlycoSuiteDB; P07911; --
Genew; HGNC:12559; UMOD.
MIM; 191845; --
InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR01881; EGF_Ca.
DR InterPro; IPR001507; Endoglin/CD105.
Pfam; PF00008; EGF; 3.
PRINTS; PR00023; ZPELUCIDA.
SMART; SM00179; EGF_CA; 2.
SMART; SM00001; EGF_Like; 1.
SMART; SM00241; ZP; 1.
PROSITE; PS00682; ZP DOMAIN; 1.
PROSITE; PS00010; ASX_HYDROXYL; 2.
PROSITE; PS00022; EGF_1; FALSE_NEG.
PROSITE; PS01186; EGF_2; 3.
PROSITE; PS01187; EGF_CA; 2.
Glycoprotein; Signal; Membrane; GPI-anchor; EGF-like domain.
FT SIGNAL 1 24
FT CHAIN 25 640 UROMODULIN.
FT DOMAIN 28 64 EGF-LIKE 1.
FT DOMAIN 65 107 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 108 149 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 334 585 ZP.
FT DISULFID 32 41 BY SIMILARITY.
FT DISULFID 35 50 BY SIMILARITY.
FT DISULFID 52 63 BY SIMILARITY.
FT DISULFID 69 83 BY SIMILARITY.
FT DISULFID 77 92 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 112 126 BY SIMILARITY.
FT DISULFID 120 135 BY SIMILARITY.
FT DISULFID 137 148 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 /FTID=CAR 000178.
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 565 565 H -> D (IN REF. 2).
SQ SEQUENCE 640 AA; 69760 MW; D26A07A76353AE48 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 640;
Best Local Similarity 57.1%; Pred. No. 51;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

Db 120 CHALATC 126

RESULT 5

```
UROM_BOVIN
ID UROM_BOVIN STANDARD; PRT; 643 AA.
AC P48733;
DT .01-FEB-1996 (Rel. 33, Created)
DT .01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Uromodulin precursor (Tamm-Horsfall urinary glycoprotein) (THP).
GN UMOD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=95143938; PubMed=7531049;
RX Yu H., Papa F., Sukhatme V.P.;
RT "Bovine and rodent Tamm-Horsfall protein (THP) genes: cloning,
RT structural analysis, and promoter identification.";
RL Gene Expr. 4:63-75(1994).
CC -!- FUNCTION: NOT KNOWN. MAY PLAY A ROLE IN REGULATING THE CIRCULATING
CC ACTIVITY OF CYTOKINES AS IT BINDS TO IL-1, IL-2 AND TNF WITH HIGH
CC AFFINITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR,
CC THEN CLEAVED TO PRODUCE A SOLUBLE FORM WHICH IS SECRETED IN
CC URINE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; S75958; AAB33312.1; -.
DR HSP; P35555; 1EMN.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR001597; Endoglin/CD105.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00100; zona pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00179; EGF CA; 2.
DR SMART; SM00001; EGF_Like; 1.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF 1; FALSE_NEG.
DR PROSITE; PS01186; EGF 2; 3.
DR PROSITE; PS01187; EGF CA; 2.
DR GlycoProtein; Signal; Membrane; GPI-anchor; EGF-like domain.
FT SIGNAL 1 26
FT CHAIN 27 643
FT DOMAIN 30 66
FT DOMAIN 67 109
FT DOMAIN 110 151
FT DOMAIN 336 587
FT DISULFID 34 43
FT DISULFID 37 52
FT DISULFID 54 65
FT DISULFID 71 85
FT DISULFID 79 94
FT DISULFID 96 108
FT DISULFID 114 128
FT DISULFID 122 137
FT DISULFID 139 150
FT CARBOHYD 40 40
FT CARBOHYD 78 78
N-LINKED (GLCNAC. . .) (POTENTIAL).

us-10-105-008-44.rsp
Page 4
```

```

ID  ARGE LRPBI      STANDARD;      PRT;    277 AA.
AC  P13440;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Possible acetylornithine deacetylase (EC 3.5.1.16)
DE  (Acetylornithinase).
GN  ARGE.
OS  Leptospira biflexa.
OC  Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
OX  NCBI_TaxID=172;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Serovar Patoc;
RX  MEDLINE=89008067; PubMed=2844724;
RA  Zuerner R.L., Charon N.W.;
RT  "Nucleotide sequence analysis of a gene cloned from Leptospira
RT  biflexa serovar patoc which complements an argE defect in Escherichia
RT  coli.";
RL  J. Bacteriol. 170:4548-4554(1988).
CC  -!- CATALYTIC ACTIVITY: N2-acetyl-L-ornithine + H(2)O = acetate + L-
CC  ornithine.
CC  -!- PATHWAY: Arginine biosynthesis; fifth step.
CC  -!- SIMILARITY: RESEMBLES PORTIONS OF THE BETA' SUBUNITS OF RNA
CC  POLYMERASES FROM BACTERIA AND CHLOROPLASTS. IT WAS SUGGESTED THAT
CC  BOTH THE BETA' SUBUNIT OF RNA POLYMERASES AND L.BIFLEXA ARGE-
CC  COMPLEMENTING ACTIVITY POSSESSES N-ACETYLASE OR N-ACETYLORNITHINASE
CC  ACTIVITY. NO HOMOLOGY WITH E.COLI ARGE.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/annouce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; M22622; AAA25260.1;
DR  PIR; A31840; A31840.
DR  HSP; Q9KWU6; 1HQM.
DR  InterPro; IPR002879; RNA_pol_A2.
DR  Pfam; PF01854; RNA_pol_A2; 1.
KW  Arginine biosynthesis; Hydrolase.
FT  SIMILAR 76 / 135 TO BETA' SUBUNITS OF RNA POLYMERASES.
FT  SIMILAR 165 / 224 TO BETA' SUBUNITS OF RNA POLYMERASES.
SQ  SEQUENCE 277 AA; 31107 MW; 40434BCFA4D7ED63 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 277;
Best Local Similarity 57.1%; Pred. No. 38;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
||| :|||
DB 265 CHTLSTC 271.

RESULT 8
VN53 ROTPC      STANDARD;      PRT;    393 AA.
AC  Q00033;
DT  01-DEC-1992 (Rel. 24, Created)
DT  01-DEC-1992 (Rel. 24, Last sequence update)
DT  01-JUN-1994 (Rel. 29, Last annotation update)
DE  Nonstructural RNA-binding protein 53 (NS53) (NCVP2).
OS  Porcine rotavirus (group C / strain Cowden).
OC  Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX  NCBI_TaxID=10916;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93277387; PubMed=8389118;
RA  Bremond M., Chabanne-Vautherot D., Cohen J.;
RT  "Sequence analysis of three non structural proteins of a porcine
RT  group C (Cowden strain) rotavirus.";
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/annouce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; M22622; AAA25260.1;
DR  PIR; A31840; A31840.
DR  HSP; Q9KWU6; 1HQM.
DR  InterPro; IPR002879; RNA_pol_A2.
DR  Pfam; PF01854; RNA_pol_A2; 1.
KW  Arginine biosynthesis; Hydrolase.
FT  SIMILAR 76 / 135 TO BETA' SUBUNITS OF RNA POLYMERASES.
FT  SIMILAR 165 / 224 TO BETA' SUBUNITS OF RNA POLYMERASES.
SQ  SEQUENCE 277 AA; 31107 MW; 40434BCFA4D7ED63 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 277;
Best Local Similarity 57.1%; Pred. No. 38;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
||| :|||
DB 265 CHTLSTC 271.

RESULT 9
AT52 BOVIN      STANDARD;      PRT;    1205 AA.
ID  AT52 BOVIN
AC  P79331;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and
DE  metalloproteinase with thrombospondin motifs 2) (ADAM-TS2)
DE  (Procollagen I/II amino-propeptide processing enzyme) (Procollagen I
DE  N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (PNPI).
GN  ADAMTS2 OR NPI.
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE FROM N.A.
RT  TISSUE=Skin;
RX  MEDLINE=97225960; PubMed=9122202;
RA  Colige A., Li S.W., Sieron A.L., Nusgens B.V., Prockop D.J.,
RA  Lapiere C.M.;
RT  "cDNA cloning and expression of bovine procollagen I N-proteinase: a
RT  new member of the superfamily of zinc-metalloproteinases with binding
RT  sites for cells and other matrix components.";
RL  Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).
RN  [2]
RP  PARTIAL SEQUENCE.
RX  MEDLINE=95348096; PubMed=7622483;
RA  Colige A., Beschin A., Samyn B., Goebels Y., Van Beeumen J.,
RA  Nusgens B.V., Lapiere C.M.;
RT  "Characterization and partial amino acid sequencing of a 107-kDa
RT  procollagen I N-proteinase purified by affinity chromatography on
RT  immobilized type XIV collagen.";
RL  J. Biol. Chem. 270:16724-16730(1995).
CC  -!- FUNCTION: CLEAVES THE PROPEPTIDES OF TYPE I AND II COLLAGEN PRIOR
CC  TO FIBRIL ASSEMBLY. DOES NOT ACT ON TYPE III COLLAGEN. MAY ALSO
CC  PLAY A ROLE IN DEVELOPMENT THAT IS INDEPENDANT OF ITS ROLE IN
CC  COLLAGEN BIOSYNTHESIS.
CC  -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC  alpha-1(I) at Pro-|-Gln and of alpha-1(II) and alpha-2(I) chains
CC  at Ala-|-Gln.
CC  -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
```

```

Arch. Virol. 130:85-92(1993).
-!- FUNCTION: THIS NONSTRUCTURAL PROTEIN IS ONLY EXPRESSED AT VERY
LOW LEVELS EARLY IN INFECTION. IT COULD HAVE SOME ROLE IN THE
SELECTION OF RNA SEGMENTS FOR PACKAGING (BY SIMILARITY).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/annouce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; X60546; CAA43036.1;
DR  PIR; A48357; A48357.
DR  PIR; S16051; S16051.
KW  Nonstructural protein; RNA-binding; Zinc-finger.
FT  ZN FING 39 68 POTENTIAL.
SQ  SEQUENCE 393 AA; 46438 MW; DB7FAD9AE2E3636F CRC64;

Query Match 76.2%; Score 32; DB 1; Length 393;
Best Local Similarity 57.1%; Pred. No. 51;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
||| :|||
DB 273 CHTLSTC 279

RESULT 9
AT52 BOVIN      STANDARD;      PRT;    1205 AA.
ID  AT52 BOVIN
AC  P79331;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and
DE  metalloproteinase with thrombospondin motifs 2) (ADAM-TS2)
DE  (Procollagen I/II amino-propeptide processing enzyme) (Procollagen I
DE  N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (PNPI).
GN  ADAMTS2 OR NPI.
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE FROM N.A.
RT  TISSUE=Skin;
RX  MEDLINE=97225960; PubMed=9122202;
RA  Colige A., Li S.W., Sieron A.L., Nusgens B.V., Prockop D.J.,
RA  Lapiere C.M.;
RT  "cDNA cloning and expression of bovine procollagen I N-proteinase: a
RT  new member of the superfamily of zinc-metalloproteinases with binding
RT  sites for cells and other matrix components.";
RL  Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).
RN  [2]
RP  PARTIAL SEQUENCE.
RX  MEDLINE=95348096; PubMed=7622483;
RA  Colige A., Beschin A., Samyn B., Goebels Y., Van Beeumen J.,
RA  Nusgens B.V., Lapiere C.M.;
RT  "Characterization and partial amino acid sequencing of a 107-kDa
RT  procollagen I N-proteinase purified by affinity chromatography on
RT  immobilized type XIV collagen.";
RL  J. Biol. Chem. 270:16724-16730(1995).
CC  -!- FUNCTION: CLEAVES THE PROPEPTIDES OF TYPE I AND II COLLAGEN PRIOR
CC  TO FIBRIL ASSEMBLY. DOES NOT ACT ON TYPE III COLLAGEN. MAY ALSO
CC  PLAY A ROLE IN DEVELOPMENT THAT IS INDEPENDANT OF ITS ROLE IN
CC  COLLAGEN BIOSYNTHESIS.
CC  -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC  alpha-1(I) at Pro-|-Gln and of alpha-1(II) and alpha-2(I) chains
CC  at Ala-|-Gln.
CC  -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
```

EMBL; X96389; CAA65253.1; --
MEROPS; M12.301; --
InterPro; IPR001762; Disintegrin.
InterPro; IPR002870; Pep_M12B_propep.
InterPro; IPR001590; Reprolysin.
InterPro; IPR000884; TSP1.
InterPro; IPR000130; Zn_MTpeptdse.
Pfam; PF00090; tsp_1; 4;
Pfam; PF01421; Reprolysin; 1.
Pfam; PF01562; Pep_M12B_propep; 1.
SMART; SM00209; TSPI; 4.
PROSITE; PS00215; ADAM_MEPRO; 1.
PROSITE; PS00092; TSPI_1.
PROSITE; PS00142; ZINC_PROTEASE; FALSE NEG.
PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Collagen degradation; Extracellular matrix; Heparin-binding.
SIGNAL 1 28
PROPEP 29 253
CHAIN 254 1205
METAL 402 402
ACT SITE 403 403
BY SIMILARITY.
ZINC (CATALYTIC) (POTENTIAL).
METAL 406 406
BY SIMILARITY.
ZINC (CATALYTIC) (BY SIMILARITY).
DOMAIN 412 412
ZINC (CATALYTIC) (BY SIMILARITY).
DOMAIN 474 554
DISINTEGRIN-LIKE.
DOMAIN 555 611
TSP TYPE-1 1.
DOMAIN 612 716
CYS-RICH.
DOMAIN 717 845
SPACER.
DOMAIN 846 905
TSP TYPE-1 2.
DOMAIN 906 968
TSP TYPE-1 3.
DOMAIN 969 1024
TSP TYPE-1 4.
SITE 685 687
CELL ATTACHMENT SITE (POTENTIAL).
DOMAIN 1 35
POLY-ALA.
DOMAIN 177 180
POLY-GLU.
CARBOHYD 104 104
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 245 245
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 942 942
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 943 943
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 987 987
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 1025 1025
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 1092 1092
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 1139 1139
N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 1144 1144
N-LINKED (GLCNAC. .) (POTENTIAL).

```

SQ  SEQUENCE  1205 AA;  133887 MW;  7B5B232A45320371 CRC64;

Query Match      76.2%;  Score 32;  DB 1;  Length 1205;
Best Local Similarity 71.4%;  Pred. No. 1.3e+02;
Matches 5;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

OY  1 CHAVGSC 7
    || || ||
DB  380 CHPVRS C 386

RESULT 10
ATS3_HUMAN
ID  ATS3_HUMAN  STANDARD;  PRT;  1205 AA.
AC  O15072; Q9BXZ8;
DT  16-OCT-2001 (Rel. 40, Created)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  ADAMTS-3 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE  with thrombospondin motifs 3) (ADAM-TS 3) (ADAM-TS3) (Procollagen II
DE  amino-propeptide processing enzyme) (Procollagen II N-proteinase) (PC
DE  II-NP).
GN  ADAMTS3 OR KIAA0366.
OS  Homo sapiens (Human).
OC  Bukaryoka; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
[1]
RN  SEQUENCE OF 1-227 FROM N.A.
RP  MEDLINE=21402912; PubMed=11408482;
RX  Fernandes R.J., Hirohata S., Engle J.M., Colige A., Cohn D.H.,
RA  Eyre D.R., Apte S.S.;
RT  "Procollagen II amino propeptide processing by ADAMTS-3. Insights on
RT  dermatopapillaxis."
RL  J. Biol. Chem. 276:31502-31509(2001).
[2]
RN  SEQUENCE OF 5-1205 FROM N.A.
RP  TISSUE=Brain;
RC  MEDLINE=97349984; PubMed=9205841;
RX  Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA  Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT  "Prediction of the coding sequences of unidentified human genes. VII.
RT  The complete sequences of 100 new cDNA clones from brain which can
RT  code for large proteins in vitro."
RL  DNA Res. 4:141-150(1997).
CC  -1- FUNCTION: Cleaves the propeptides of type II collagen prior to
CC  fibril assembly. Does not act on types I and III collagens.
CC  -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC  matrix (by similarity).
CC  -1- TISSUE SPECIFICITY: Found in cartilage and skin.
CC  -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC  FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC  -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC  SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC  -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC  -1- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC  -1- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS4.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/annouce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF247668; AAK28400.1; -.
DR  EMBL; AB002364; BAA20821.1; -.
DR  MEROPS; M12.220; -.
DR  Genew; HGNC:219; ADAMTS3.
DR  MIM; 605011; -.

```


Hypothetical 39.2 kDa protein in EGT2-KRE1 intergenic region.

YNL326C OR N0325.

OS Saccharomycetes cerevisiae (Baker's Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes.

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

OX [1]

RP SEQUENCE FROM N.A.

RP STRAIN=S288C / FY1676;

RC MEDLINE=95373280; PubMed=7645347;

RX MEDLINE=95373280; PubMed=7645347;

RA Maftahi M., Nicaud J.-M., Levesque H., Gaillardin C.;

RT "Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV

RT identifies the RPD3, PAS8 and KRE1 loci, five new open reading

RT frames.;"

RT Yeast 11:567-572(1995).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: TO C.ELEGANS ZK757.1 AND S.POMBE SPBC2F12.15C.

CC -1- SIMILARITY: CONTAINS 1 DHHC-TYPE ZINC FINGER.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z46259; CAA863372.1; -;

DR EMBL; Z71602; CAA96258.1; -;

DR SGD; S0005270; YNL326C.

DR InterPro; IPR001594; Znf_DHHC.

DR Pfam; PF01529; zf-DHHC; 1.

DR ProDom; PD003041; Znf_DHHC; 1.

DR PROSITE; PS50216; ZF_DHHC; 1.

KM Hypothetical protein; Transmembrane; Zinc-finger.

FT TRANSMEM 38 58 POTENTIAL.

FT TRANSMEM 148 168 POTENTIAL.

FT TRANSMEM 189 209 POTENTIAL.

FT ZN FING 104 154 DHHC-TYPE.

FT SEQUENCE 336 AA; 39183 MW; D36BF37DEC5D3983 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 336;

Best Local Similarity 71.4%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CHAVSSC 7

DB 117 CHHCSSC 123

RESULT 14

NOF2_DROME

ID_NOF2_DROME STANDARD; PRT; 403 AA.

AC P15297;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 46 kDa protein in NOP-FB transposable element.

GN NOP.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

OX [1]

RP SEQUENCE FROM N.A.

RP STRAIN=Oregon-R;

RC MEDLINE=89356666; PubMed=2548860;

RA Templeton N.S., Potter S.S.;

RT "Complete foldback transposable elements encode a novel protein found

RT in Drosophila melanogaster.;"

RL EMBO J. 8:1887-1894(1989).

CC -1- FUNCTION: NOT KNOWN.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: TO C-TERMINAL OF NOP-FB ELEMENT PROTEIN IN
 CC STRAIN TE146(Z).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X15469; CAA33497.1; -;
 CC PIR; S07825; S07825.
 CC FlyBase; FBgn0044029; NOPORF.
 CC Nuclear protein; Transposable element.
 CC
 CC SEQUENCE 403 AA; 46519 MW; 52606117BF64A164 CRC64;
 SQ
 Query Match 73.8%; Score 31; DB 1; Length 403;
 Best Local Similarity 57.1%; Pred. No. 80;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 |||: ||
 Db 336 CHRICSC 342
 RESULT 15
 ID TSA4 GIALA STANDARD; PRT; 713 AA.
 AC P21849;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major surface-labeled trophozoite antigen 417 precursor.
 GN TSA 417.
 OS Giardia lamblia (Giardia intestinalis).
 CC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
 CC NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 30957 / WB;
 RX MEDLINE=90280395; PubMed=2352929;
 RA Gillin F.D., Hagblom P., Harwood J., Aley S.B., Reiner D.S.,
 RA McCaffery M., So M., Guiney D.G.;
 RT "Isolation and expression of the gene for a major surface protein of
 RT Giardia lamblia."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4463-4467(1990).
 RN [2]
 RP SEQUENCE OF 480-620 FROM N.A.
 RC STRAIN=AD-1;
 RX MEDLINE=93314970; PubMed=8325510;
 RA Ey P.L., Mayrhofer G.;
 RT "Two genes encoding homologous 70-kDa surface proteins are present
 RT within individual trophozoites of the binucleate protozoan parasite
 RT Giardia intestinalis."
 RL Gene 129:257-262(1993).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ANCHORED TO THE
 CC PLASMA MEMBRANE.
 CC -!- DOMAIN: CONTAINS 29 REPEATS OF THE CXXC MOTIF.
 CC -!- SIMILARITY: BELONGS TO THE GIARDIA VARIANT SURFACE PROTEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M33641; AAA02688.1; -;
 CC EMBL; M97488; AAA02581.1; -;
 CC PIR; A35502; A35502.

DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR005127; Giardia_VSP.
 DR Pfam; PF03302; VSP; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00001; EGF-like; 1.
 DR SMART; SM00261; FU; 3.
 KW Signal; Antigen; Glycoprotein; Transmembrane; Repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 713 MAJOR SURFACE-LABELED TROPHOZOITE
 FT ANTIGEN 417.
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM
 FT DOMAIN 18 679
 FT TRANSMEM 680 708
 FT DOMAIN 709 713
 FT CARBOHYD 289 289
 FT CARBOHYD 676 676
 FT VARIANT 582 582
 FT VARIANT 606 606
 FT SEQUENCE 713 AA; 72510 MW; 9AD7195843DE5601 CRC64;
 SQ
 Query Match 73.8%; Score 31; DB 1; Length 713;
 Best Local Similarity 71.4%; Pred. No. 1; 3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 |||: ||
 Db 331 CEAVSNC 337
 RESULT 16
 ID SAL3 HUMAN STANDARD; PRT; 1300 AA.
 AC Q9BXA9; OSUGH1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sal-like protein 3 (Zinc finger protein SALL3) (HSALL3).
 GN SALL3.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RA Gooding R., Angelicheva D., Blechschmidt K., Swoboda K., Molnar M.,
 RA Tournev I., Kalydjieva L.;
 RT "Exclusion of HSALL3 and refinement of the region for the CCFDN
 RT gene."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 29-1300 FROM N.A. AND ALTERNATIVE SPLICING.
 RX MEDLINE=20079154; PubMed=10610715;
 RA Kohlhaase J., Hausmann S., Stojmenovic G., Dixkens C., Bink K.,
 RA Schulz-Schaeffer W., Altmann M., Engel W.;
 RT "SALL3, a new member of the human spalt-like gene family, maps to
 RT 18q23."
 RL Genomics 62:216-222(1999).
 CC -!- FUNCTION: Probable transcription factor.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1, 2, 3 (shown here) and 4; are
 CC produced by alternative splicing. Isoforms 1 and 2 lack two zinc
 CC finger domains and are the major isoforms.
 CC -!- TISSUE SPECIFICITY: Widely expressed in adult with highest levels
 CC in heart. Expressed in fetal brain (in neurons of hippocampus,
 CC cortex, mediodorsal and ventrolateral thalamic nuclei, putamen,
 CC cerebellum and brainstem).
 CC -!- DEVELOPMENTAL STAGE: In fetal brain of the 24th gestational week.
 CC -!- SIMILARITY: BELONGS TO THE SAL FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; AF347021; AAK18311.1; -.
CC EMBL; AJ007421; CAB65124.1; -.
CC HSP; P07248; IARE.
CC Genew; HGNC:10527; SALL3.
CC MIM; 605079; -.
CC DR InterPro; IPR000822; Znf_C2H2.
CC DR Pfam; PF00096; zf-C2H2; 10.
CC DR SMART; SM00355; Znf_C2H2_1; 9.
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 8.
CC DR Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger;
KW Metal-binding; Repeat; Alternative splicing.
FT ZN_FING 420 442
FT ZN_FING 448 470
FT ZN_FING 679 701
FT ZN_FING 707 729
FT ZN_FING 739 761
FT ZN_FING 977 999
FT ZN_FING 1005 1027
FT ZN_FING 1113 1135
FT ZN_FING 1141 1163
FT DOMAIN 198 201
FT DOMAIN 213 216
FT DOMAIN 897 902
FT VARSPIC 1 133
FT VARSPIC 973 1044
FT CONFLICT 235 235
FT CONFLICT 593 593
FT CONFLICT 787 787
FT CONFLICT 797 802
FT CONFLICT 808 808
FT CONFLICT 1138 1138
FT CONFLICT 1141 1141
FT CONFLICT 1141 1141
FT SEQUENCE 1300 AA; 135371 MW; 82CF3BDCB6D59150 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 1300;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 684 CHRVLSC 690

RESULT 17
SALL_MOUSE STANDARD; PRT; 1322 AA.
AC Q9ER74; Q920R5;
AT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sal-like protein 1 (zinc finger protein Spalt-3) (Sal-3) (MSal-3).
GN SALL1 OR SALL3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.;
RC STRAIN=129/Ola;
RX MEDLINE=20422467; PubMed=10965108;
RA Buck A.; Archangelo L.; Dixkens C.; Kohlbase J.;
RT "Molecular cloning, chromosomal localization, and expression of the
RL murine SALL1 ortholog Sall-1.";
RN Cytogenet. Cell Genet. 89:150-153(2000).
[2]
SEQUENCE FROM N.A.;
RX MEDLINE=21543523; PubMed=11688560;

```

```

RA Nishinakamura R., Matsumoto Y., Nakao K., Nakamura K., Sato A.,
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Scully S., Lacey D.L.,
RA Katsuki M., Asashima M., Yokota T.;
RT "Murine homolog of SALL1 is essential for ureteric bud invasion in
RL kidney development.";
RL Development 128:3105-3115(2001).
RN [3]
RP FUNCTION, AND ASSOCIATION WITH HDAC.
RX MEDLINE=21964044; PubMed=11836251;
RA Kiefer S.M., McDill B.W., Yang J., Rauchman M.;
RT "Murine sall1 represses transcription by recruiting a histone
RL deacetylase complex.";
RL J. Biol. Chem. 277:14869-14876(2002).
CC -!- FUNCTION: Transcriptional repressor involved in organogenesis.
CC Essential for ureteric bud invasion in kidney development.
CC Homozygous deletion of SALL1 results in an incomplete ureteric bud
CC outgrowth, a failure of tubule formation in the mesenchyme and an
CC apoptosis of the mesenchyme.
CC -!- SUBUNIT: Interacts with HDAC1, HDAC2, RBP4, RBP7, MTA1 and MTA2.
CC Probably associates with NURD histone deacetylase complex (HDAC).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed in the metanephric mesenchyme
CC surrounding ureteric bud.
CC -!- SIMILARITY: BELONGS TO THE SALL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).


EMBL; AJ271914; CAC09602.1; -.
EMBL; AJ271915; CAC09602.1; JOINED.
EMBL; AB051409; BAB5673.1; -.
HSSP; P15822; LBBO.
DR MGD; MGI:1889585; Sall1.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 10.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 9.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 9.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;
KW Zinc-finger; Metal-binding; Repeat.
FT ZN_FING 450 472
FT ZN_FING 478 500
FT ZN_FING 705 727
FT ZN_FING 733 755
FT ZN_FING 765 787
FT ZN_FING 1000 1022
FT ZN_FING 1028 1050
FT ZN_FING 1133 1155
FT ZN_FING 1161 1183
FT DOMAIN 133 139
FT DOMAIN 237 240
FT DOMAIN 1143 1146
FT CONFLICT 165 165
FT CONFLICT 167 167
FT CONFLICT 1271 1271
FT SEQUENCE 1322 AA; 140229 MW; DF4FEF7FEA0B9F5C CRC64;

Query Match 73.8%; Score 31; DB 1; Length 1322;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 710 CHRVLSC 716


```

RESULT 18
SAL3 MOUSE STANDARD; PRT; 1323 AA.
AC Q62255;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sal-like protein 3 (Spalt-like protein 3) (MSal) (Fragment).
GN SALL3 OR SAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain, and Embryo;
RC MEDLINE=96391179; PubMed=8798152;
RA Ott T., Kaestner K.H., Monaghan A.P., Schuetz G.;
RT "The mouse homolog of the region specific homeotic gene spalt of
RT Drosophila is expressed in the developing nervous system and in
RT mesoderm-derived structures.";
RL Mech. Dev. 56:117-128 (1996).
CC -!- FUNCTION: Probable transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing. Isoform 2 lacks two zinc finger
CC domains (6 and 7) and is the major isoform.
CC -!- TISSUE SPECIFICITY: In adult brain, testis and kidney. In lower
CC levels also in adult ovaries and embryonic stem cells. In embryo
CC in developing neuroectoderm of brain, inner ear and spinal chord.
CC Also weakly and transiently expressed in embryonic branchial
CC arches, notochord, limb buds and heart.
CC -!- DEVELOPMENTAL STAGE: During embryogenesis detected from 7 dpc
CC onward in tissues derived from mesoderm and ectoderm.
CC -!- SIMILARITY: BELONGS TO THE SAL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X97581; CRA66196.1; -
CC HSSP; P07248; IARE.
CC MGI; MGI:109295; Sall3.
CC InterPro; IPR000822; Znf C2H2.
CC Pfam; PF00096; Zf-C2H2; 10.
CC SMART; SM00355; Znf C2H2; 9.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 8.
CC Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger;
CC Metal-binding; Repeat; Alternative splicing.
CC
CC NON TER 1 1
CC ZN_FING 398 420 C2H2-TYPE.
CC ZN_FING 426 448 C2H2-TYPE.
CC ZN_FING 663 685 C2H2-TYPE.
CC ZN_FING 691 713 C2H2-TYPE.
CC ZN_FING 723 745 C2H2-TYPE.
CC ZN_FING 968 990 C2H2-TYPE.
CC ZN_FING 996 1018 C2H2-TYPE.
CC ZN_FING 1104 1126 C2H2-TYPE.
CC ZN_FING 1132 1154 C2H2-TYPE.
CC DOMAIN 119 123 POLY-PRO.
CC DOMAIN 188 191 POLY-GLN.
CC DOMAIN 345 348 POLY-SER.
CC DOMAIN 881 886 POLY-SER.
CC VARSPLIC 964 1035 MISSING (IN ISOFORM 2).
CC SEQUENCE 1323 AA; 139070 MW; 86D83DEAF2CED0D CRC64;
v Match 73.8%; Score 31; DB 1; Length 1323;

Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CHAVSSC 7
DB 668 CHRVLSC 674
RESULT 19
SALL1 HUMAN STANDARD; PRT; 1324 AA.
ID SALL1 HUMAN
AC Q9NSC2; Q9NSC3; Q99881; Q9PIR0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sal-like protein 1 (Zinc finger protein SALL1) (Spalt-like
DE transcription factor 1) (HSall).
GN SALL1 OR SALL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.; DISEASE, AND VARIANTS SER-150 DEL; SER-150 INS AND
RP GLY-159.
RX MEDLINE=99138656; PubMed=9973281;
RA Kohlhaase J., Taschner P.E.M., Surfeind P., Pasche B., Newman B.,
RA Blanck C., Breuning M.H., ten Kate L.P., Maaswinkel-Mooy P.,
RA Mitulla B., Seidel J., Kirkpatrick S.J., Pauli R.M., Wargowski D.S.,
RA Devriendt K., Proesmans W., Gabrielli O., Coppa G.V.,
RA Wesby-van Swaay E., Trembath R.C., Schinzel A.A., Reardon W.,
RA Seemanova E., Engel W.;
RT "Molecular analysis of SALL1 mutations in Townes-Brocks Syndrome.";
RL Am. J. Hum. Genet. 64:435-445 (1999).
RN [2]
RP SEQUENCE OF 1-26 FROM N.A., DISEASE, AND VARIANTS SER-164 DEL AND
RP GLU-1265.
RX MEDLINE=20004537; PubMed=10533063;
RA Marlin S., Blanchard S., Lacombe D., Denoyelle F., Alessandri J.-L.,
RA Calzolari E., Drouin-Garraud V., Ferraz F.G., Fourmaintraux A.,
RA Philip N., Toubanc J.E., Petit C.;
RT "Townes-Brocks syndrome: Detection of a SALL1 mutation hot spot and
RT evidence for a position effect in one patient.";
RL Hum. Mutat. 14:377-386 (1999).
RN [3]
RP SEQUENCE OF 26-1324 FROM N.A.
RX MEDLINE=97131507; PubMed=8975705;
RA Kohlhaase J., Schuh R., Dove G., Kuehnlein R.P., Jaekle H.,
RA Schroeder B., Schulz-Schaeffer W., Kretzschmar H.A., Koehler A.,
RA Mueller U., Raab-Vetter M., Burkhardt E., Engel W., Stick R.;
RT "Isolation, characterization, and organ-specific expression of two
RT novel human zinc finger genes related to the Drosophila gene spalt.";
RL Genomics 38:291-298 (1996).
RN [4]
RP SEQUENCE OF 313-345 FROM N.A. AND DISEASE.
RX MEDLINE=98085876; PubMed=9425907;
RA Kohlhaase J., Wischermann A., Reichenbach H., Froster U., Engel W.;
RT "Mutations in the SALL1 putative transcription factor gene cause
RT Townes-Brocks syndrome.";
RL Nat. Genet. 18:82-83 (1998).
RN [5]
RP DISEASE.
RX MEDLINE=20381976; PubMed=10928856;
RA Engels S., Kohlhaase J., McGaughran J.;
RT "A SALL1 mutation causes a branchio-oto-renal syndrome-like
RT phenotype.";
RL J. Med. Genet. 37:458-460 (2000).
CC -!- FUNCTION: Transcriptional repressor involved in organogenesis (By
CC similarity).
CC -!- SUBUNIT: Interacts with HDAC2, HDAC3, RBBP4, RBBP7, MTA1 and MTA2
CC (By similarity). Probably associates with Nurd histone deacetylase
CC complex (HDAC).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
 DB 4 CHITKSC 10

RESULT 21

IR10_HCMVA
 ID IR10_HCMVA STANDARD; PRT; 171 AA.
 AC P16808;
 DT 01-AUG-1990 (Rel. 15, Last Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Hypothetical protein IR10 precursor (TRL10).
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]

SEQUENCE FROM N.A.
 RP MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Hornell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 cytomegalovirus strain AD169";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X17403; CAA35300.1; -
 DR PIR; S09903; S09903.
 DR PIR; S09759; S09759.
 KW Hypothetical protein; Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 171 HYPOTHETICAL PROTEIN IRL10/IRL10.
 FT TRANSMEM 80 100 POTENTIAL.
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 171 AA; 19034 MW; 2C2E8AD869419B86 CRC64;

Query Match / 71.4%; Score 30; DB 1; Length 171;
 Best Local Similarity 57.1%; Pred. No. 59;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
 DB 60 CHAMCKC 66

RESULT 22

DHSB_ECOLI
 ID DHSB_ECOLI STANDARD; PRT; 238 AA.
 AC P07014;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1).
 GN SDHB OR B0724.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=85046453; PubMed=6388571;
 RA Darlison M.G., Guest J.R.;
 RT "Nucleotide sequence encoding the iron-sulphur protein subunit of the
 RL succinate dehydrogenase of Escherichia coli."; Biochem. J. 223:507-517(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Ikenoma T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 12.7-28.0 min region on the linkage map";
 RL DNA Res. 3:137-155(1996).
 RN [4]
 RP SEQUENCE OF 231-238 FROM N.A.
 RX MEDLINE=84236168; PubMed=6376123;
 RA Darlison M.G., Spencer M.E., Guest J.R.;
 RT "Nucleotide sequence of the suca gene encoding the 2-oxoglutarate
 dehydrogenase of Escherichia coli K12";
 RL Eur. J. Biochem. 141:351-359(1984).
 RN [5]
 RP SEQUENCE OF 1-11.
 RC STRAIN=K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 in the genome of Escherichia coli K-12";
 RL Electrophoresis 18:1359-1313(1997).
 CC -!- FUNCTION: TWO DISTINCT, MEMBRANE-BOUND, FAD-CONTAINING ENZYMES ARE
 CC RESPONSIBLE FOR THE CATALYSIS OF FUMARATE AND SUCCINATE
 CC INTERCONVERSION; THE FUMARATE REDUCTASE IS USED IN ANAEROBIC
 CC GROWTH, AND THE SUCCINATE DEHYDROGENASE IS USED IN AEROBIC GROWTH.
 CC -!- CATALYTIC ACTIVITY: Succinate + acceptor = fumarate + reduced
 CC acceptor.
 CC -!- COFACTOR: BINDS THREE DIFFERENT IRON-SULFUR CLUSTERS: A 2FE-2S,
 CC A 3FE-4S AND A 4FE-4S.
 CC -!- PATHWAY: Tricarboxylic acid cycle.
 CC -!- SUBUNIT: PART OF AN ENZYME COMPLEX CONTAINING FOUR SUBUNITS: A
 CC FLAVOPROTEIN, AN IRON-SULFUR, CYTOCHROME B-556, AND AN HYDROPHOBIC
 CC ANCHOR PROTEIN.
 CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
 CC 'PLANT-TYPE' 2FE-2S AND 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J01619; AAA23896.1; -
 DR EMBL; X01070; CAA25534.1; -
 DR EMBL; AB000175; AAC73818.1; -
 DR EMBL; D90711; BAA35391.1; -
 DR EMBL; X00661; CAA25279.1; -

```
DR PIR: A28837; DECSI.
DR HSP: P00364; 1FUM.
DR Ecogene; EGI0932; sdhB.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR004489; DhsB.
DR InterPro; IPR001041; Ferredoxin.
DR Pfam; PF00111; fer2; 1.
DR TIGRfams; TIGR00384; dhsB; 1.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 1.
KW Tricarboxylic acid cycle; Iron-sulfur; Oxidoreductase.
FT METAL 55 55 IRON-SULFUR S1 (2FE-2S) (BY SIMILARITY).
FT METAL 60 60 IRON-SULFUR S1 (2FE-2S) (BY SIMILARITY).
FT METAL 75 75 IRON-SULFUR S1 (2FE-2S) (BY SIMILARITY).
FT METAL 149 149 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
FT METAL 152 152 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
FT METAL 155 155 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
FT METAL 159 159 IRON-SULFUR S3 (3FE-4S) (BY SIMILARITY).
FT METAL 206 206 IRON-SULFUR S3 (3FE-4S) (BY SIMILARITY).
FT METAL 212 212 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
FT METAL 216 216 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 238 AA; 26770 MW; 226F60C55FAC35A CRC64;

Query Match 71.4%; Score 30; DB 1; Length 238;
Best Local Similarity 42.9%; Pred. No. 78;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 206 CHSINMC 212

RESULT 23
DHSB_SALTY
ID DHSB_SALTY STANDARD; PRT; 238 AA.
AC Q8ZOU2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1).
GN SDHB OR STM0735.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -!- FUNCTION: TWO DISTINCT, MEMBRANE-BOUND, FAD-CONTAINING ENZYMES ARE
CC RESPONSIBLE FOR THE CATALYSIS OF FUMARATE AND SUCCINATE
CC INTERCONVERSION; THE FUMARATE REDUCTASE IS USED IN ANAEROBIC
CC GROWTH, AND THE SUCCINATE DEHYDROGENASE IS USED IN AEROBIC GROWTH
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Succinate + acceptor = fumarate + reduced
CC acceptor.
CC -!- COFACTOR: BINDS THREE DIFFERENT IRON-SULFUR CLUSTERS: A 2FE-2S,
CC A 3FE-4S AND A 4FE-4S.
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SUBUNIT: PART OF AN ENZYME COMPLEX CONTAINING FOUR SUBUNITS: A
CC FLAVOPROTEIN, AN IRON-SULFUR, CYTOCHROME B-556, AND AN HYDROPHOBIC
CC ANCHOR PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'PLANT-TYPE' 2FE-2S AND 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE008730; AAL19679.1; ALT_INIT.
CC StyGene; SG7777; sdhB.
CC InterPro; IPR001450; 4Fe4S_ferredoxin.
CC InterPro; IPR004489; DhsB.
CC InterPro; IPR001041; Ferredoxin.
CC Pfam; PF00111; fer2; 1.
CC TIGRfams; TIGR00384; dhsB; 1.
CC PROSITE; PS00198; 4Fe4S_FERREDOXIN; 1.
KW Tricarboxylic acid cycle; Iron-sulfur; Oxidoreductase.
FT METAL 55 55 IRON-SULFUR S1 (2FE-2S) (BY SIMILARITY).
FT METAL 60 60 IRON-SULFUR S1 (2FE-2S) (BY SIMILARITY).
FT METAL 75 75 IRON-SULFUR S1 (2FE-2S) (BY SIMILARITY).
FT METAL 149 149 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
FT METAL 152 152 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
FT METAL 155 155 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
FT METAL 159 159 IRON-SULFUR S3 (3FE-4S) (BY SIMILARITY).
FT METAL 206 206 IRON-SULFUR S3 (3FE-4S) (BY SIMILARITY).
FT METAL 212 212 IRON-SULFUR S3 (3FE-4S) (BY SIMILARITY).
FT METAL 216 216 IRON-SULFUR S2 (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 238 AA; 26733 MW; 7628F9F9F95611F3 CRC64;

Query Match 71.4%; Score 30; DB 1; Length 238;
Best Local Similarity 42.9%; Pred. No. 78;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
DB 206 CHSINMC 212

RESULT 24
AMPM_ECOLI
ID AMPM_ECOLI STANDARD; PRT; 264 AA.
AC P07506;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methionine aminopeptidase (EC 3.4.11.18) (MAP) (Peptidase M).
GN MAP OR B0168 OR Z0178 OR ECS0170.
OS Escherichia coli.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87109068; PubMed=3027045;
RX Ben-Bassat A., Bauer K., Chang S.-Y., Myambo K., Boosman A., Chang S.;
RT "Processing of the initiation methionine from proteins: properties of
RT the Escherichia coli methionine aminopeptidase and its gene
RT structure."
RL J. Bacteriol. 169:751-757(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RX MEDLINE=94261430; PubMed=8202364;
RX Fujita N., Mori H., Yura T., Ishihama A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 2.4-4.1 min (110,917-193,643 bp) region."
RL Nucleic Acids Res. 22:1637-1639(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
```


RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Submitted 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Laskari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouisis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
RA Iida T., Takai H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=93229487; PubMed=8471602;
RA Roderick S.L., Mathews B.W.;
RT "Structure of the cobalt-dependent methionine aminopeptidase from
RT *Escherichia coli*: a new type of proteolytic enzyme.";
RL Biochemistry 32:3907-3912(1993).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=99316170; PubMed=10387007;
RA Lowther W.T., Orville A.M., Madden D.T., Lim S., Rich D.H.,
RA Mathews B.W.;
RT "Escherichia coli methionine aminopeptidase: implications of
RT crystallographic analyses of the native, mutant, and inhibited
RT enzymes for the mechanism of catalysis.";
RL Biochemistry 38:7678-7688(1999).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX MEDLINE=20025411; PubMed=10555963;
RA Lowther W.T., Zhang Y., Sampson P.B., Honek J.F., Matthews B.W.;
RT "Insights into the mechanism of *Escherichia coli* methionine
RT aminopeptidase from the structural analysis of reaction products and
RT phosphorus-based transition-state analogues.";
RL Biochemistry 38:14810-14819(1999).
CC -1- FUNCTION: REMOVES THE AMINO-TERMINAL METHIONINE FROM NASCENT
CC PROTEINS.
CC -1- CATALYTIC ACTIVITY: Release of N-terminal amino acids,
CC preferentially methionine, from peptides and arylamides.
CC -1- COFACTOR: COBALT; BINDS 2 IONS PER SUBUNIT, SODIUM; BINDS 1 ION
CC PER SUBUNIT.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15106; AAA24112.1; -;
DR EMBL; D26562; BAA05612.1; -;
DR EMBL; AB000126; AAC73279.1; -;
DR EMBL; U70214; AAB08597.1; -;
DR EMBL; AB005192; AAG54470.1; -;
DR EMBL; AP002550; BAB33593.1; -;
DR PIR; A27761; DPECM.
DR PIR; S45233; S45233.
DR PDB; 1MAT; 31-JAN-94.
DR PDB; 2MAT; 18-JUN-99.
DR PDB; 3MAT; 18-JUN-99.
DR PDB; 4MAT; 18-JUN-99.
DR PDB; 1C21; 17-NOV-99.
DR PDB; 1C22; 17-NOV-99.
DR PDB; 1C23; 17-NOV-99.
DR PDB; 1C24; 17-NOV-99.
DR PDB; 1C27; 17-NOV-99.
DR MEROPS; M24.001; -;
DR SWISS-2DPAGE; P07906; COLI.
DR ECODBASE; F029.7; 6TH EDITION.
DR EcoGene; EG10570; map.
DR InterPro; IPR002467; MAP_1.
DR InterPro; IPR001714; Methamino_Ptase.
DR InterPro; IPR000994; Peptidase M24.
DR Pfam; PF00557; Peptidase M24; 1.
DR PRINTS; PR00599; MAPEPTIDASE.
DR TIGRFAMs; TIGR00500; met_pdase_1; 1.
DR PROSITE; PS00680; MAP_1; 1.
DR Hydrolase; Aminopeptidase; Cobalt; 3D-structure; Complete proteome.
FT METAL 97 97 COBALT 2.
FT METAL 108 108 COBALT 1 AND 2.
FT METAL 171 171 COBALT 1.
FT METAL 204 204 COBALT 1.
FT METAL 235 235 COBALT 1 AND 2.
FT HELIX 8 28
FT TURN 29 31
FT TURN 34 35
FT STRAND 37 37
FT HELIX 38 50
FT TURN 51 52
FT STRAND 56 57
FT TURN 60 61
FT HELIX 62 64
FT STRAND 70 73
FT TURN 74 75
FT STRAND 76 78
FT TURN 84 85
FT STRAND 87 87
FT STRAND 93 102
FT TURN 103 104
FT STRAND 105 114
FT HELIX 120 139
FT TURN 142 143
FT STRAND 145 145
FT HELIX 146 158
FT TURN 159 161
FT STRAND 163 164
FT STRAND 170 172
FT STRAND 179 183
FT TURN 189 190
FT STRAND 194 194
FT TURN 197 198
FT STRAND 200 203
FT STRAND 206 208
FT STRAND 214 216
FT TURN 218 219
FT STRAND 223 225
FT STRAND 231 233
FT STRAND 235 241

FT TURN 242 243
FT STRAND 244 247
FT TURN 248 249
FT TURN 252 253
FT STRAND 259 261
SQ SEQUENCE 264 AA; F2D0B57715A67851 CRC64;

Query Match 71.4%; Score 10; DB 1; Length 264;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
Db 54 HAVSAC 59

RESULT 25

UL53 HCMVA STANDARD; PRT; 376 AA.
AC PL6794;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Protein UL53 (HERF2 protein).
GN UL53.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL31,
CC EHV-1 29, EBV BFLF2, HCMV UL53, AND VZV 27.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X17403; CAA35412.1; --
DR PIR; S09816; Q9SEW2.
DR InterPro; IPR003868; Herpes UL31-like.
DR Pfam; PF02718; Herpes_UL31; 1.
SQ SEQUENCE 376 AA; 42312 MW; BF4D5F8DE2B8B9 CRC64;

Query Match 71.4%; Score 10; DB 1; Length 376;
Best Local Similarity 83.3%; Pred. No. 11e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVSS 6
Db 245 CHAVSA 250

Search completed: June 20, 2003, 20:53:06
Job time : 10.4318 secs

20-JAN-1999;
99US-0234395.

Sequence 8 AA;

Cell adhesion modulation; CAM; synaptic stability; cadherin; cadherin-mediated adhesion; drug delivery; cell adhesion; tumor wound healing; neurite outgrowth; cyclic; circular.

OS Unidentified.
 XX WO200053632-A1.
 XX 14-SEP-2000.
 XX 07-MAR-2000; 2000WO-CA00222.
 XX 09-MAR-1999; 99US-0265107.
 XX (UYMC-) UNIV MCGILL.
 XX Blaschuk OW, Gour BJ;
 XX WPI; 2000-594308/56.
 XX Stimulating beta-catenin mediated gene expression, cellular
 PT differentiation and hair growth, involves contacting cells with
 PT modulating agent capable of inhibiting interaction between alpha and
 PT beta catenin
 XX Disclosure; Page 6; 77pp; English.
 XX The present invention is concerned with methods of modulating the amount
 CC of free beta-catenin in the cell, and methods of stimulating the
 CC expression of genes involved in cellular differentiation, the
 CC transcription of which is under the control of beta-catenin. The
 CC peptides given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351
 CC can be used as modulating agents which interrupt the interaction between
 CC alpha and beta catenin, causing increased levels of the latter and
 CC stimulating the activation of beta-catenin mediated transcription. This
 CC can be used to stimulate cell differentiation, which can then be used to
 CC promote hair growth and skin exfoliation. This latter is particularly
 CC useful in the improvement of photodamaged skin and to minimise wrinkles.
 CC The modulating peptide can also be used to reduce hearing loss resulting
 CC from inner ear disorders such as hyperacusis and tinnitus.
 XX Sequence 7 AA;
 SQ Query Match 78.6%; Score 33; DB 21; Length 7;
 Best Local Similarity 71.4%; Pred. No. 7.8e+05;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CHAVSSC 7
 Db 1 CHAVVNC 7
 RESULT 21
 AAB27344
 ID AAB27344 standard; Peptide; 7 AA.
 XX AAB27344;
 XX 15-FEB-2001 (first entry)
 XX Beta-catenin derived cyclic peptide SEQ ID NO: 32.
 XX Beta-catenin; cadherin-mediated intercellular adhesion;
 KW cell differentiation; modulating agent; hair loss; skin exfoliation;
 KW internalisation moiety; flanking sequence; transcription; hearing loss;
 KW cyclic.
 XX Unidentified.
 XX Key Location/Qualifiers
 FH Disulfide-bond 1..7
 FT /note= "the disulfide bond cyclises the molecule"
 XX WO200053632-A1.
 XX 14-SEP-2000.
 XX 07-MAR-2000; 2000WO-CA00222.
 XX 09-MAR-1999; 99US-0265107.
 XX (UYMC-) UNIV MCGILL.
 XX Blaschuk OW, Gour BJ;
 XX WPI; 2000-594308/56.
 XX Stimulating beta-catenin mediated gene expression, cellular
 PT differentiation and hair growth, involves contacting cells with
 PT modulating agent capable of inhibiting interaction between alpha and
 PT beta catenin
 XX Disclosure; Page 7; 77pp; English.
 XX The present invention is concerned with methods of modulating the amount
 CC of free beta-catenin in the cell, and methods of stimulating the
 CC expression of genes involved in cellular differentiation, the
 CC transcription of which is under the control of beta-catenin. The
 CC peptides given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351
 CC can be used as modulating agents which interrupt the interaction between
 CC alpha and beta catenin, causing increased levels of the latter and
 CC stimulating the activation of beta-catenin mediated transcription. This
 CC can be used to stimulate cell differentiation, which can then be used to
 CC promote hair growth and skin exfoliation. This latter is particularly
 CC useful in the improvement of photodamaged skin and to minimise wrinkles.
 CC The modulating peptide can also be used to reduce hearing loss resulting
 CC from inner ear disorders such as hyperacusis and tinnitus.
 XX Sequence 7 AA;
 SQ Query Match 78.6%; Score 33; DB 21; Length 7;
 Best Local Similarity 71.4%; Pred. No. 7.8e+05;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CHAVSSC 7
 Db 1 CHAVVNC 7
 RESULT 22
 AAW23984
 ID AAW23984 standard; peptide; 8 AA.
 XX AAW23984;
 XX 23-JUL-1998 (first entry)
 XX Cadherin-mediated cell adhesion modulating peptide 6.
 XX Cadherin-mediated; cell adhesion; drug delivery; treatment; cancer;
 KW angiogenesis; skin grafting; neurite growth; veterinary medicine;
 KW wound healing; tumour; metastasis; carcinoma; leukaemia; modulation;
 KW organ transplant; neurological disease.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "N-terminal acetyl"
 FT Disulfide-bond 1
 FT /note= "disulphide bridge"
 FT Modified-site 8
 FT /note= "C-terminal amide"
 FT Disulfide-bond 8
 FT /note= "disulphide bridge"
 XX WO9802452-A2.
 XX 22-JAN-1998.
 XX

DE Peptide inhibiting interaction of alpha-catenin and beta-catenin.
 XX Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;
 KW cadherin-mediated function; demyelinating neurological disease;
 KW multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;
 KW central nervous system; apoptosis induction; cadherin-expression cell;
 KW pregnancy prevention; vasopermeability; synaptic stability; diabetes;
 KW rheumatoid arthritis; allergic response; learning; memory;
 KW antibody-mediated graft rejection; cyclic.

OS Synthetic.

XX WO9845319-A2.

PN 15-OCT-1998.

PD 14-APR-1998; 98WO-CA00322.

PP 10-APR-1997; 97US-0043361.

PR (UYMC-) UNIV MCGILL.

PA Blaschuk OW, Gour BJ;

PI WPI; 1999-024009/02.

DR New catenin modulating agents - comprising peptides having a

PT sequence HAV or analogues or antibodies, used for modulating

PT cadherin-mediated functions

XX Claim 12; Page 78; 106pp; English.

CC The present sequence represents a peptide which is capable of inhibiting
 CC an interaction between alpha-catenin and beta-catenin. The peptide is
 CC used in modulating agents that are used for modulating cadherin-mediated
 CC functions. They can be used for disrupting interaction between
 CC alpha-catenin and beta-catenin in a cell, inhibiting cell adhesion, e.g.
 CC between epithelial cells, endothelial cells, neural cells, tumour cells
 CC and lymphocytes, for treating a demyelinating neurological disease, e.g.
 CC multiple sclerosis, for reducing unwanted cellular adhesion in a mammal,
 CC for enhancing the delivery of a drug through the skin of a mammal, for
 CC enhancing the delivery of a drug to a tumour in a mammal, for treating
 CC cancer in a mammal, for inhibiting angiogenesis in a mammal, for
 CC inducing apoptosis in a cadherin-expression cell, for modulating the
 CC immune system of a mammal, for preventing pregnancy in a mammal, for
 CC increasing vasopermeability in a mammal, or for inhibiting synaptic
 CC stability in a mammal. In particular they can be used for treating
 CC diabetes, rheumatoid arthritis, allergic responses, antibody-mediated
 CC graft rejection or for stimulating learning and memory.

XX Sequence 7 AA;

Query Match 78.6%; Score 33; DB 20; Length 7;
 Best Local Similarity 71.4%; Pred. No. 7.8e+05;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7

DB 1 CHAVVNC 7

RESULT 19

AAW91019

ID AAW91019 standard; peptide; 7 AA.

XX AAW91019;

AC AAW91019;

XX 24-MAR-1999 (first entry)

DE Peptide inhibiting interaction of alpha-catenin and beta-catenin.

XX Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;

KW cadherin-mediated function; demyelinating neurological disease;
 KW multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;
 KW central nervous system; apoptosis induction; cadherin-expression cell;
 KW pregnancy prevention; vasopermeability; synaptic stability; diabetes;
 KW rheumatoid arthritis; allergic response; learning; memory;
 KW antibody-mediated graft rejection.

XX Synthetic.

OS WO9845319-A2.

PN 15-OCT-1998.

PP 14-APR-1998; 98WO-CA00322.

XX 10-APR-1997; 97US-0043361.

XX (UYMC-) UNIV MCGILL.

XX Blaschuk OW, Gour BJ;

XX WPI; 1999-024009/02.

XX New catenin modulating agents - comprising peptides having a

PT sequence HAV or analogues or antibodies, used for modulating

PT cadherin-mediated functions

XX Claim 10; Page 77; 106pp; English.

CC The present sequence represents a peptide which is capable of inhibiting
 CC an interaction between alpha-catenin and beta-catenin. The peptide is
 CC used in modulating agents that are used for modulating cadherin-mediated
 CC functions. They can be used for disrupting interaction between
 CC alpha-catenin and beta-catenin in a cell, inhibiting cell adhesion, e.g.
 CC between epithelial cells, endothelial cells, neural cells, tumour cells
 CC and lymphocytes, for treating a demyelinating neurological disease, e.g.
 CC multiple sclerosis, for reducing unwanted cellular adhesion in a mammal,
 CC for enhancing the delivery of a drug through the skin of a mammal, for
 CC enhancing the delivery of a drug to a tumour in a mammal, for treating
 CC cancer in a mammal, for inhibiting angiogenesis in a mammal, for
 CC enhancing drug delivery to the central nervous system of a mammal, for
 CC inducing apoptosis in a cadherin-expression cell, for modulating the
 CC immune system of a mammal, for preventing pregnancy in a mammal, for
 CC increasing vasopermeability in a mammal, or for inhibiting synaptic
 CC stability in a mammal. In particular they can be used for treating
 CC diabetes, rheumatoid arthritis, allergic responses, antibody-mediated
 CC graft rejection or for stimulating learning and memory.

XX Sequence 7 AA;

Query Match 78.6%; Score 33; DB 20; Length 7;
 Best Local Similarity 71.4%; Pred. No. 7.8e+05;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7

DB 1 CHAVVNC 7

RESULT 20

AAW91019

ID AAW91019 standard; peptide; 7 AA.

XX AAW91019;

AC AAW91019;

XX 15-FEB-2001 (first entry)

DE Beta-catenin derived linear peptide SEQ ID NO: 20.

XX Beta-catenin; cadherin-mediated intercellular adhesion;

XX cell differentiation; modulating agent; hair loss; skin exfoliation;

XX internalisation moiety; flanking sequence; transcription; hearing loss.

XX AAY65418;
AC
XX
DT 01-FEB-2000 (first entry)
DE
XX
KW Human 5' EST related polypeptide SEQ ID NO:1579.
DE
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification.
XX
XX OS Homo sapiens.
XX
XX WO9953051-A2.
PN
XX
XX 21-OCT-1999.
PD
XX
XX 09-APR-1999; 99WO-IB00712.
PF
XX
XX 09-APR-1998; 98US-0057719.
PR
XX 28-APR-1998; 98US-0069047.
PR
XX
XX (GEST) GENSET.
PA
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX
XX WPI; 2000-038446/63.
DR
XX N-PSDB; AAZ43032.
DR
XX
XX Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
PT
XX
XX Claim 3; Page 827; 837pp; English.
XX
XX AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag. (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65438 represent the EST-related proteins corresponding to AAZ42265 to
CC AAZ43082. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AAZ42249 to AAZ42264 and AAY64644 to AAY64650 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 108 AA;
SQ
Query Match 81.0%; Score 34; DB 21; Length 108;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CHAVSSC 7
DB 14 CHRVSAC 20
RESULT 17
ABG10977
ID ABG10977 standard; Protein; 175 AA.
XX
XX ABG10977;
XX

DT 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #10968.
DE
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
KW
XX OS Homo sapiens.
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR
XX 23-AUG-2000; 2000US-0649167.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI; 2001-639362/73.
DR
XX N-PSDB; RAS75164.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 41336; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 175 AA;
SQ
Query Match 81.0%; Score 34; DB 22; Length 175;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CHAVSSC 7
DB 111 CHTVSGC 117
RESULT 18
AAW91031
ID AAW91031 standard; peptide; 7 AA.
XX
XX AAW91031;
AC
XX
XX 24-MAR-1999 (first entry)
DT
XX

CC at: ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 419 AA;
Query Match 85.7%; Score 36; DB 22; Length 419;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CHAVSSC 7
| | | | |
Db 162 CHSLSSC 168
RESULT 14
AAM40445
ID AAM40445 standard; Protein; 1302 AA.
AC AAM40445;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5376.
XX
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren P, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59601.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 5376; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM3642-AAI42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC (HYSE-)

CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1302 AA;
Query Match 83.3%; Score 35; DB 22; Length 1302;
Best Local Similarity 71.4%; Pred. No. 8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CHAVSSC 7
| | | | |
Db 695 CHALSQC 701
RESULT 15
AAM84690
ID AAM84690 standard; Protein; 41 AA.
XX
AC AAM84690;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:12283.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226888.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.

KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

XX Homo sapiens.

XX W0200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -

XX Claim 27; SEQ ID No 31585; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.

XX Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 343 AA;

Query Match 90.5%; Score 38; DB 23; Length 343;

Best Local Similarity 85.7%; Pred. No. 72;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7

DB 80 CHVSSC 86

RESULT 13

ABGL7301

ID ABGL7301 standard; Protein; 419 AA.

XX AC ABGL7301;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #17292.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN W0200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSBQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS81488.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 47660; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX SQ Sequence 343 AA;

Query Match 90.5%; Score 38; DB 22; Length 343;

Best Local Similarity 85.7%; Pred. No. 72;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CHAVSSC 7

DB 80 CHVSSC 86

RESULT 10

AAM72106

ID AAM72106 standard; Protein; 343 AA.

XX AC

XX AAM72106;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32412.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX PS Example 4; SEQ ID NO: 32412; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.

XX SQ Sequence 343 AA;

Query Match 90.5%; Score 38; DB 22; Length 343;

Best Local Similarity 85.7%; Pred. No. 72;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CHAVSSC 7

DB 80 CHVSSC 86

RESULT 11

AAM32368

ID AAM32368 standard; Protein; 343 AA.

XX AC AAM32368;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #6405 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX PS Claim 27; SEQ ID NO 32637; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
 CC of these probes. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.

XX SQ Sequence 343 AA;

Query Match 90.5%; Score 38; DB 22; Length 343;

Best Local Similarity 85.7%; Pred. No. 72;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CHAVSSC 7

DB 80 CHVSSC 86

RESULT 12

ABG41920

ID ABG41920 standard; Peptide; 343 AA.

XX AC ABG41920;

XX DT 19-AUG-2002 (first entry)

XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 31585.

XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;

PR 06-APR-1999; 99US-0287150.
PR 13-MAY-1999; 99US-0311021.
XX (GEM) GENETICS INST INC.
PA Wong GG, Clark HF, Fechtel K, Agostino MJ;
XX PI
XX PT
XX DR WPI; 2000-053095/04.
XX DR N-PSDB; AA252528.
XX
XX PT Novel polynucleotides and proteins having biological activities which
XX PT make them suitable for treating, preventing or ameliorating medical
XX PT conditions in humans or animals.
XX
XX PS Claim 117; Page 654-655; 730pp; English.
XX
XX CC The present invention describes human secreted proteins encoded by
XX CC polynucleotides obtained from adult testes, foetal brain, adult brain,
XX CC brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
XX CC cDNA libraries. The polynucleotides and proteins are predicted to have
XX CC biological activities which would make them suitable for treating,
XX CC preventing or ameliorating medical conditions in humans and animals.
XX CC Suggested activities include nutritional activity, cytokine and cell
XX CC proliferation/differentiation activity, immune stimulating (e.g. as
XX CC vaccines) or suppressing activity, haematopoiesis regulating activity,
XX CC tissue growth activity, activin/inhibin activity, chemotactic/
XX CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
XX CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
XX CC suppressor activity, and tumour inhibition activity. The polynucleotides
XX CC are also stated to be useful for gene therapy. Therapeutic compositions
XX CC are also presently valuable for veterinary applications. AA252475 to
XX CC AA252581 encode human secreted proteins, and AA273390 to AA273500
XX CC represent human secreted proteins, given in the present invention.
XX
XX SQ Sequence. 109 AA;
XX

Query Match 90.5%; Score 38; DB 21; Length 109;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CHAVSSC 7
DB 100 CHGVSSC 106
II IIII
RESULT 8
ABB38887
ID ABB38887 standard; Peptide; 343 AA.
AC ABB38887;
XX
XX DT 04-FEB-2002 (first entry)
XX
XX DE Peptide #6393 encoded by human foetal liver single exon probe.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX OS Homo sapiens.
XX
XX PN WO200157277-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00669.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX

(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-483447/52.
Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human foetal liver.
Claim 27; SEQ ID NO 31522; 639pp + sequence listing; English.
The invention relates to a single exon nucleic acid probe for
measuring human gene expression in a sample derived from human foetal
liver. The single exon nucleic acid probes may be used for predicting,
measuring and displaying gene expression in samples derived from human
foetal liver. The present sequence is a peptide encoded by a single exon
nucleic acid probe of the invention.
Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 343 AA;
Query Match 90.5%; Score 38; DB 22; Length 343;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CHAVSSC 7
DB 80 CHVSSC 86
II IIII
RESULT 9
AAM59538
ID AAM59538 standard; Protein; 343 AA.
XX
XX AC AAM59538;
XX
XX DT 05-NOV-2001 (first entry)
XX
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31643.
XX KW Human; brain expressed exon; gene expression analysis; probe;
XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX KW epilepsy; cancer.
XX
XX OS Homo sapiens.
XX
XX PN WO200157275-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00667.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX

(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-483446/52.
Single exon nucleic acid probes for analyzing gene expression in human
brains.
Example 4; SEQ ID NO: 31643; 650pp + Sequence Listing; English.
XX

XX PD 26-JUL-2001.
 XX PF 24-JAN-2001; 2001WO-US02508.
 XX PR 24-JAN-2000; 2000US-0491078.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 XX PT Hu Z;
 XX DR WPI; 2001-549899/61.
 XX PT Cell adhesion modulating agent used for enhancing delivery of drug to
 XX PT tumor comprises imidazole compounds -
 XX PS Examples; Fig 3A-I; 436pp; English.
 XX CC The invention relates to cell adhesion modulating agents that comprise
 XX CC imidazole compounds of specified formulae that are peptidomimetics of
 XX CC cyclic peptides. The peptidomimetics have a structure similar to that of
 XX CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 XX CC The agents are used for modulating classical cadherin mediated
 XX CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 XX CC delivery of a drug to a tumor, inhibiting development of cancer,
 XX CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 XX CC wound healing, modulating the immune system, increasing vasopermeability,
 XX CC treating demyelinating disease, facilitating migration of an N-cadherin
 XX CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 XX CC neurite outgrowth, and treating spinal cord injuries and macular
 XX CC degeneration. The present sequence represents a cyclic control peptide.
 XX SQ Sequence 7 AA;
 Query Match 90.5%; Score 38; DB 22; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CHAVSSC 7
 DB 1 CHGVSSC 7
 RESULT 6
 ID AAM47512 standard; peptide; 7 AA.
 AC AAM47512;
 XX DT 12-FEB-2002 (first entry)
 XX DE Cyclic control peptide #19.
 XX KW Cadherin; cytostatic; gynecological; endometriosis;
 XX KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 XX KW cyclic.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 XX FT Modified-site 1
 XX FT /note= "Linked to residue 7 to form a cyclic peptide,
 XX FT N-terminal acetyl"
 XX FT Modified-site 7
 XX FT /note= "Linked to residue 1 to form a cyclic peptide,
 XX FT C-terminal amide"
 XX WO200177146-A2.
 XX PN 18-OCT-2001.
 XX PD 09-APR-2001; 2001WO-US11669.

XX PR 07-APR-2000; 2000US-0544782.
 XX PA (UTMC-) UNIV MCGILL.
 XX PI Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 XX DR WPI; 2002-049129/06.
 XX PT Modulating endothelial cell adhesion for inhibiting development of
 XX PT endometriosis, increasing blood flow to tumor in a mammal, by
 XX PT contacting cell with a cyclic peptide having cadherin cell adhesion
 XX PT recognition sequence -
 XX PS Disclosure; Fig 3; 139pp; English.
 XX CC The present invention relates to a method for modulating endothelial cell
 XX CC adhesion. The method comprises contacting an endothelial cell with a
 XX CC modulating peptide comprising a cadherin cell adhesion recognition
 XX CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 XX CC for inhibiting angiogenesis, increasing vasopermeability, increasing
 XX CC blood flow to a tumour, disrupting neovasculation and inhibiting the
 XX CC development of endometriosis in a mammal. The modulating peptide reduces
 XX CC unwanted endothelial adhesion occurring between tumour cells, tumour
 XX CC cells and normal cells, normal cells as a result of surgery, injury,
 XX CC chemotherapy, disease and inflammation. The present sequence is a peptide
 XX CC used as a control for illustrating the present invention.
 XX SQ Sequence 7 AA;
 Query Match 90.5%; Score 38; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CHAVSSC 7
 DB 1 CHGVSSC 7
 RESULT 7
 ID AAY73443 standard; Protein; 109 AA.
 AC AAY73443;
 XX DT 29-FEB-2000 (first entry)
 XX DE Human secreted protein clone yb187_1 protein sequence SEQ ID NO:108.
 XX KW Human; secreted protein; immunostimulatory; haemostatic; cytokine;
 XX KW proliferative; differentiative; chemotactic; chemokinetic; vaccine;
 XX KW thrombolytic; antiinflammatory; cytostatic; immunosuppressive;
 XX KW gene therapy.
 XX OS Homo sapiens.
 XX PN WO9958642-A2.
 XX PD 18-NOV-1999.
 XX XX 14-MAY-1999; 99WO-US10843.
 XX PR 14-MAY-1998; 98US-0085472.
 XX PR 17-AUG-1998; 98US-0096824.
 XX PR 11-SEP-1998; 98US-0099843.
 XX PR 11-SEP-1998; 98US-0099950.
 XX PR 15-SEP-1998; 98US-0100424.
 XX PR 29-SEP-1998; 98US-0102329.
 XX PR 09-OCT-1998; 98US-0103615.
 XX PR 11-DEC-1998; 98US-0111799.
 XX PR 14-DEC-1998; 98US-0112159.
 XX PR 31-DEC-1998; 98US-0114415.
 XX PR 10-FEB-1999; 99US-0248059.

XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "putative N-terminal acetylation or
 FT FT alkoxybenzoylation"
 FT Modified-site 7 /note= "putative C-terminal amide or ester"
 FT FT
 FN WO200153331-A2.
 XX 26-JUL-2001.
 XX 24-JAN-2001; 2001WO-US02508.
 XX 24-JAN-2000; 2000US-0491078.
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX Gour BJ, Blaschuk OW, Ali A, Ni F, Chen Z, Michaud SD, Wang S;
 PI Hu Z;
 XX WPI; 2001-549899/61.
 XX Cell adhesion modulating agent used for enhancing delivery of drug to
 FT tumor comprises imidazole compounds -
 XX
 XX Examples; Fig 3A-I; 436pp; English.
 XX The invention relates to cell adhesion modulating agents that comprise
 CC imidazole compounds of specified formulae that are peptidomimetics of
 CC cyclic peptides. The peptidomimetics have a structure similar to that of
 CC a cyclic peptide that comprises a cadherin cell recognition sequence HAV.
 CC The agents are used for modulating classical cadherin mediated
 CC intercellular adhesion, reducing unwanted cellular adhesion, enhancing
 CC delivery of a drug to a tumor, inhibiting development of cancer,
 CC inhibiting angiogenesis, enhancing drug delivery to the CNS, enhancing
 CC wound healing, modulating the immune system, increasing vasopermeability,
 CC treating demyelinating disease, facilitating migration of an N-cadherin
 CC expressing cell on astrocytes, inhibiting synaptic stability, modulating
 CC neurite outgrowth, and treating spinal cord injuries and macular
 CC degeneration. The present sequence represents a cyclic peptide with
 CC cadherin cell adhesion recognition (CAR) sequence.
 XX Sequence 7 AA;
 SQ Query Match 100.0%; Score 42; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CHAVSSC 7
 Db 1 CHAVSSC 7
 RESULT 4
 AAM47511
 ID AAM47511 standard; peptide; 7 AA.
 XX AC AAM47511;
 XX 12-FEB-2002 (first entry)
 DT Cyclic peptide endothelial cell adhesion modulator #27.
 XX Cadherin; cytostatic; gynecological; endometriosis;
 KW endothelial cell adhesion; angiogenesis; vasopermeability; tumour;
 KW cyclic.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "Linked to residue.7 to form a cyclic peptide,
 FT FT

FT Modified-site 7 N-terminal acetyl"
 FT FT /note= "Linked to residue 1 to form a cyclic peptide,
 FT FT C-terminal amide"
 XX WO200177146-A2.
 XX 18-OCT-2001.
 XX 09-APR-2001; 2001WO-US11669.
 XX 07-APR-2000; 2000US-0544782.
 XX (UTMC-) UNIV MCGILL.
 XX Blaschuk OW, Gour BJ, Farookhi R, Ali A;
 XX WPI; 2002-049129/06.
 XX Modulating endothelial cell adhesion for inhibiting development of
 FT endometriosis, increasing blood flow to tumor in a mammal, by
 FT contacting cell with a cyclic peptide having cadherin cell adhesion
 FT recognition sequence -
 XX Example 4; Page 55; 139pp; English.
 XX The present invention relates to a method for modulating endothelial cell
 CC adhesion. The method comprises contacting an endothelial cell with a
 CC modulating peptide comprising a cadherin cell adhesion recognition
 CC sequence (His Ala Val) within a cyclic peptide ring. The method is useful
 CC for inhibiting angiogenesis, increasing vasopermeability, increasing the
 CC blood flow to a tumour, disrupting neovasculation and inhibiting the
 CC development of endometriosis in a mammal. The modulating peptide reduces
 CC unwanted endothelial adhesion occurring between tumour cells, tumour
 CC cells and normal cells, normal cells as a result of surgery, injury, such
 CC chemotherapy, disease and inflammation. The present sequence is one
 CC modulating peptide.
 XX Sequence 7 AA;
 SQ Query Match 100.0%; Score 42; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CHAVSSC 7
 Db 1 CHAVSSC 7
 RESULT 5
 AAG65405
 ID AAG65405 standard; peptide; 7 AA.
 XX AC AAG65405;
 XX 30-NOV-2001 (first entry)
 DT Cyclic control peptide sequence.
 XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 KW cytostatic; vulnary; immunomodulator; vasotropic; neuroprotective;
 KW cerebroprotective; muscular; cyclic.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "putative N-terminal acetylation or
 FT FT alkoxybenzoylation"
 FT Modified-site 7 /note= "putative C-terminal amide or ester"
 FT FT
 XX WO200153331-A2.

XX AAY17154;
 XX 20-SEP-1999 (first entry)
 XX Cadherin-mediated adhesion modulating cyclic peptide.
 XX Cell adhesion modulation; CAM; synaptic stability; cadherin;
 XX cadherin-mediated adhesion; drug delivery; cell adhesion; tumour;
 XX wound healing; neurite outgrowth; cyclic; circular.
 XX Synthetic.
 XX WO9933875-A1.
 XX 08-JUL-1999.
 XX 23-DEC-1998; 98WO-CA01207.
 XX 23-DEC-1997; 97US-0996679.
 XX (UYMC-) UNIV MCGILL.
 XX Blaschuk OW, Gour BJ;
 XX WPI; 1999-430231/36.
 XX Cyclic peptide cell adhesion modulating agents, useful for
 XX modulating synaptic stability
 XX Claim 11; Page 65; 144pp; English.
 XX The invention provides cyclic peptide cell adhesion modulating (CAM)
 XX agents that comprises a His-Ala-Val recognition sequence. Also provided
 XX is a method for inhibiting synaptic stability in a mammal that comprises
 XX administering to a mammal a therapeutically effective amount of a CAM
 XX agent that inhibits cadherin-mediated adhesion, where the agent comprises
 XX a cyclic peptide having a peptide ring, and where the sequence His-Ala-
 XX Val is present within the peptide ring. The cyclic peptides are cell
 XX adhesion modulating agents that inhibit cadherin-mediated adhesion. They
 XX can be used in a method for inhibiting synaptic stability in mammals. The
 XX agents can be used to treat diseases or other conditions characterized by
 XX undesirable cell adhesion or to facilitate drug delivery to a specific
 XX tissue or tumour. Alternatively the agents may be used to enhance cell
 XX adhesion (e.g. to supplement or replace stitches or to facilitate wound
 XX healing) or to enhance or direct neurite outgrowth.
 XX SQ Sequence 7 AA;
 Query Match 100.0%; Score 42; DB 20; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 Db 1 CHAVSSC 7
 RESULT 2
 ID AAY73802
 XX AAY73802 standard; Peptide; 7 AA.
 XX AAY73802;
 XX 07-MAR-2000 (first entry)
 XX N-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:308.
 XX Diagnosis; cancer; OB-cadherin; N-cadherin; E-cadherin; CAR; cyclic;
 XX cell adhesion recognition sequence; leukaemia; metastasis;
 XX prostate cancer; breast cancer; ovarian cancer; carcinoma.
 XX Synthetic.
 OS

OS Homo sapiens.
 XX WO9957565-A2.
 XX 11-NOV-1999.
 XX 05-MAY-1999; 99WO-CA00362.
 XX 05-MAY-1998; 98US-0073040.
 XX 06-NOV-1998; 98US-0187859.
 XX 20-JAN-1999; 99US-0234395.
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX Blaschuk OW, Gour BJ, Byers S;
 XX WPI; 2000-062165/05.
 XX Detecting expression of OB-cadherin or N-cadherin used for diagnosing
 XX and evaluating cancer
 XX Disclosure; Page 22; 56pp; English.
 XX The present invention describes methods which have been developed for
 XX detecting and evaluating cancer that are based on the finding that
 XX OB-cadherin and N-cadherin are expressed by metastatic carcinoma cells,
 XX but not by highly differentiated, poorly invasive carcinomas. A method
 XX for determining the presence or absence of a cancer in a patient
 XX comprises: (1) contacting a biological sample from the patient with an
 XX binding agent that specifically binds to OB- or N-cadherin OR with an
 XX oligonucleotide that hybridises to a polynucleotide that encodes OB- or
 XX N-cadherin; and (2) detecting in the sample an amount of polypeptide
 XX that binds to the binding agent OR the amount of polynucleotide that
 XX hybridises to the oligonucleotide, relative to a predetermined cut-off
 XX value, and determining the presence or absence of cancer in the patient
 XX from this. The methods from the present invention can be used to
 XX determine the metastatic potential of a cancer. The methods may be used
 XX to detect a metastatic cancer in a patient, to monitor progression of a
 XX cancer, or to evaluate the metastatic potential of a cancer. Cancers
 XX which may be evaluated using the methods include leukemia, prostate
 XX cancer, breast cancer and ovarian cancer. AAY73501 to AAY73813 represent
 XX peptide sequences used in the exemplification of the present invention.
 XX Specifically, AAY73503 to AAY73789 represent OB-cadherin cell adhesion
 XX recognition (CAR) peptides and AAY73790 to AAY73808 represent N-cadherin
 XX CAR peptides. AAZ47906 to AAZ47915 represent PCR primers used in examples
 XX from the present invention.
 XX SQ Sequence 7 AA;
 Query Match 100.0%; Score 42; DB 21; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CHAVSSC 7
 Db 1 CHAVSSC 7
 RESULT 3
 ID AAG65404
 XX AAG65404 standard; peptide; 7 AA.
 XX AAG65404;
 XX 30-NOV-2001 (first entry)
 XX Cyclic peptide with cadherin CAR sequence.
 XX Cell adhesion; imidazole; peptidomimetic; cadherin; angiogenesis; CAR;
 XX cytosolic; vulnery; immunomodulator; vasotropic; neuroprotective;
 XX cerebroprotective; muscular; cyclic.
 XX Synthetic.
 OS

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:32:15 ; Search time 44.7045 Seconds
(without alignments)
20.865 Million cell updates/sec

Title: US-10-105-008-44
Perfect score: 42
Sequence: 1 CHAVSSC 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : A_Geneseq_101002:*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	42	100.0	7	20	AA117154
2	42	100.0	7	21	AA117154
3	42	100.0	7	22	AA117154
4	42	100.0	7	23	AA117154
5	38	90.5	7	23	AA117154
6	38	90.5	7	23	AA117154
7	38	90.5	109	21	AA117154
8	38	90.5	343	22	AA117154
9	38	90.5	343	22	AA117154
10	38	90.5	343	22	AA117154

11	38	90.5	343	22	AA117154	Peptide #6405 enco
12	38	90.5	343	23	ABG41920	Human peptide enco
13	35	85.7	419	22	ABG17301	Novel human diagno
14	35	83.3	1302	22	AA104445	Human polypeptide
15	34	81.0	41	22	AA104460	Human immune/haema
16	34	81.0	108	21	AA165418	Human 5' EST relat
17	34	81.0	175	22	ABG10977	Novel human diagno
18	33	78.6	7	20	AA110131	Peptide inhibiting
19	33	78.6	7	20	AA110131	Peptide inhibiting
20	33	78.6	7	21	AA127332	Beta-catenin deriv
21	33	78.6	7	21	AA127332	Beta-catenin deriv
22	33	78.6	8	19	AA123984	Cadherin-mediated
23	33	78.6	8	20	AA117118	Cadherin-mediated
24	33	78.6	8	21	AA173801	N-cadherin cell ad
25	33	78.6	8	22	AA165402	Cyclic peptide wit
26	33	78.6	8	23	AA147509	Cyclic peptide end
27	33	78.6	9	19	AA123992	Cadherin-mediated
28	33	78.6	9	19	AA123992	Cadherin-mediated
29	33	78.6	9	20	AA117131	Peptide Seq ID No:
30	33	78.6	9	20	AA117132	Peptide Seq ID No:
31	33	78.6	9	22	AA165431	Cyclic peptide wit
32	33	78.6	9	22	AA165432	Cyclic peptide wit
33	33	78.6	9	23	AA147538	Peptide, SEQ ID 71
34	33	78.6	9	23	AA147539	Peptide, SEQ ID 72
35	33	78.6	10	19	AA123991	Cadherin-mediated
36	33	78.6	10	20	AA117130	Peptide Seq ID No:
37	33	78.6	10	22	AA165430	Cyclic peptide wit
38	33	78.6	10	23	AA147537	Peptide, SEQ ID 70
39	33	78.6	54	21	AA151975	Human secreted pro
40	33	78.6	72	22	AA185427	Human immune/haema
41	33	78.6	77	23	ABP04741	Human ORFX protein
42	33	78.6	101	16	AA181439	Hepatitis GB virus
43	33	78.6	101	21	AA1809016	Hepatitis GB virus
44	33	78.6	120	21	AA1809016	Human secreted pro
45	33	78.6	140	20	AA159723	Secreted protein 6
46	33	78.6	167	22	AA183465	Protonibacterium
47	33	78.6	204	23	AB181467	Murine YACH protei
48	33	78.6	531	22	AB183988	Drosophila melanog
49	33	78.6	540	22	AB171680	Drosophila melanog
50	33	78.6	570	22	AB163990	Drosophila melanog
51	33	78.6	1528	22	AB184007	Drosophila melanog
52	33	78.6	3164	16	AA194345	Hepatitis GB virus
53	32	76.2	7	22	AA165410	Cyclic peptide wit
54	32	76.2	7	23	AA147517	Cyclic peptide end
55	32	76.2	11	22	AA176500	Peptide which bind
56	32	76.2	39	20	AA107777	Human secreted pro
57	32	76.2	50	22	AA148331	Protonibacterium
58	32	76.2	65	22	AB143419	Peptide #10925 enc
59	32	76.2	65	22	AB163990	Protein #8389 enco
60	32	76.2	65	22	AA164344	Human brain expres
61	32	76.2	65	22	AA177170	Human bone marrow
62	32	76.2	65	22	AA121104	Peptide #7538 enco
63	32	76.2	65	22	AA137304	Peptide #11341 enc
64	32	76.2	65	22	AB146186	Human peptide enco
65	32	76.2	80	22	AA132321	Novel human secret
66	32	76.2	153	22	AA161158	Protonibacterium
67	32	76.2	215	22	AA129389	Human G protein-co
68	32	76.2	215	23	AB160677	Novel G protein co
69	32	76.2	225	22	AB131343	Novel human diagno
70	32	76.2	330	22	AB102170	Novel human diagno
71	32	76.2	566	19	AA147029	Human N-proteinase
72	32	76.2	969	21	AA153900	Amino acid sequenc
73	32	76.2	1201	21	AA121223	Haemophilus somnus
74	32	76.2	1201	21	AA121254	Human metalloprote
75	32	76.2	1203	22	AA150004	Bovine metalloprot

ALIGNMENTS

RESULT 1
AA117154
ID AA117154 standard; peptide; 7 AA.

RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
 RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
 evolution.";
 RL Science 0:0-0(1999).
 DR EMBL; AF091245; AAD08922.1; -;
 DR HSSP; P08151; 2GLI.
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF00096; Zf-C2H2; 5.
 DR SMART; SM00355; Znf_C2H2; 5.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
 KW DNA-binding; Metal-binding; Zinc-finger.
 FT NON TER 466 466
 SQ SEQUENCE 466 AA; 50036 MW; D5F7E39641E12A1 CRC64;

Query Match 82.5%; Score 33; DB 5; Length 466;

Best Local Similarity 83.3%; Pred. No. 79;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6

Db 312 CHWDC 317

RESULT 24

Q9E226 PRELIMINARY; PRT; 527 AA.

AC Q9E226;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Orf119-like protein.

OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;

OC Nucleopolyhedrovirus.

OX NCBI_TaxID=10468;

RN [1]

RP SEQUENCE FROM N.A.

RA Robertson A.P.S.;

RT "Genetic organization of Helicoverpa zea nuclear polyhedrosis virus in

the region of EcoRI U,D,L,A and O.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF275264; AAG17375.1; -;

SQ SEQUENCE 527 AA; 59935 MW; E1DF20478A77574C CRC64;

Query Match 82.5%; Score 33; DB 12; Length 527;

Best Local Similarity 83.3%; Pred. No. 88;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6;

Db 69 CHAVDC 74

RESULT 25

Q9VVD2 PRELIMINARY; PRT; 715 AA.

AC Q9VVD2;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE CG7729 protein (LD28629P).

GN CG7729.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=BERKELEY;

PX MEDLINE=20196006; PubMed=10731132;

RA Adams W.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balwle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoetin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003525; AAP49380.1; -;
 DR EMBL; AY061360; AAL28908.1; -;
 DR Flybase; FBgn0036688; CG7729.
 DR InterPro; IPR000299; Band_4.1.
 DR InterPro; IPR001849; PH.
 DR Pfam; PF00169; PH; 1.
 DR SMART; SM00295; B41; 1.
 DR SMART; SM00233; PH; 1.
 DR PROSITE; PSS0057; BAND_41_3; 2.
 DR PROSITE; PSS0003; PH_DOMAIN; 1.
 SQ SEQUENCE 715 AA; 81501 MW; 975E4A069643F54A CRC64;

Query Match 82.5%; Score 33; DB 5; Length 715;

Best Local Similarity 66.7%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6

Db 673 CHSADC 678

Search completed: June 20, 2003, 20:57:14

Job time : 39.6364 secs

```

Db          27 CHCLDC 32

RESULT 20
Q98K76      PRELIMINARY;      PRT; 138 AA.
AC Q98K76;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein m111601.
GN M11601.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF002997; BAB48938.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 138 AA; 14450 MW; 88D93E6688D49010 CRC64;

Query Match      82.5%; Score 33; DB 16; Length 138;
Best Local Similarity 66.7%; Pred. NO. 26;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          1 CHAVDC 6
Db          25 CHCLDC 30

RESULT 21.
Q8TYK3      PRELIMINARY;      PRT; 176 AA.
AC Q8TYK3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Uncharacterized conserved protein.
GN MK0293.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva V.A., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AF010327; AAM01510.1; --
KW Complete proteome.
SQ SEQUENCE 176 AA; 19551 MW; BFD4037D9F071C5F CRC64;

Query Match      82.5%; Score 33; DB 17; Length 176;
Best Local Similarity 83.3%; Pred. NO. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY          1 CHAVDC 6
Db          41 CHAVPC 46

RESULT 22
Q9ARGO      PRELIMINARY;      PRT; 246 AA.
AC Q9ARGO;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Osmotin-like protein (Thaumatin-like protein).
GN PR5 F23 OR TLP.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. RED MARRAS;
RA Fuchs H.C.I., Radauer C., Ebner C., Jensen-Jarolim E., Breiteneder H.,
RA Scheiner O., Hoffmann-Sommergruber K.;
RT "Characterization of the osmotin-like protein from paprika (Capsicum
RT annuum).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kim Y.S., Park J.Y., Kim K.S., Ko M.K., Chung S.J., Oh B.-J.;
RT "A thaumatin-like gene is differentially expressed in compatible and
RT incompatible interactions between pepper (Capsicum annuum) and the
RT anthracnose fungus, Colletotrichum gloeosporioides.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297410; CAC34055.2; --
DR EMBL; AF297646; AAK97184.1; --
DR HSP; F25871; IAUN.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; thaumatin; 1.
DR PRINTS; PR00347; THAUMATIN.
DR ProDom; PD001321; Thaumatin; 1.
DR SMART; SM00205; THN; 1.
DR PROSITE; PS00316; THAUMATIN; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 246 AA; 26358 MW; ASBE9288424D7637 CRC64;

Query Match      82.5%; Score 33; DB 10; Length 246;
Best Local Similarity 66.7%; Pred. NO. 44;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          1 CHAVDC 6
Db          141 CHAIQC 146

RESULT 23
Q96564      PRELIMINARY;      PRT; 466 AA.
AC Q96564;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cubitus interruptus (Fragment).
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OX NCBI_TaxID=39708;
RN [1]
RP SEQUENCE FROM N.A.
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,

```


RT genes from Phytophthora infestans.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF494014; AM18483.1; -
SQ SEQUENCE 745 AA; 82998 MW; E876636B522CB9E4 CRC64;

Query Match 85.0%; Score 34; DB 10; Length 745;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 4; Conservative 1; Mismatches *1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
DB 473 CHMDC 478

RESULT 17
Q94H19 PRELIMINARY; PRT; 1074 AA.
AC Q94H19;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Putative gag-pol polyprotein, 3'-partial (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
ON NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,
RA Utterback T.R., Relbolyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RA "Oryza sativa chromosome 3 BAC OSUNBa0077G22 genomic sequence.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084831; AAK52148.1; -
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR PROSITE; P500136; SUBTILASE_ASP; UNKNOWN_1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 1074 AA; 121171 MW; 2F25AE21F4572749 CRC64;

Query Match 85.0%; Score 34; DB 10; Length 1074;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
DB 1005 CHAIC 1010

RESULT 18
Q9QYGI PRELIMINARY; PRT; 1128 AA.
AC Q9QYGI;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Phospholipase C-L2.
GN PLC2 OR PLCE2 OR PLC-L2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050058; PubMed=10581172;

RA Otsuki M., Fukami K., Kohno T., Yokota J., Takenawa T.;
RT "Identification and characterization of a new phospholipase C-like
RL protein, PLC-L2.";
RL Biochem. Biophys. Res. Commun. 266:97-103(1999).
DR EMBL; AB033615; BAA89457.1; -
DR HSSP; P10688; 1QAS.
DR MGD; MGI:1352756; Plc12.
DR InterPro; IPR000008; C2.
DR InterPro; IPR001849; PH.
DR InterPro; IPR001192; PI_PLC.
DR InterPro; IPR000909; PI_PLC_Xdom.
DR InterPro; IPR001711; PI_PLC_Y.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00388; PI-PLC-X; 1.
DR Pfam; PF00387; PI-PLC-Y; 1.
DR PRINTS; PRO0390; PHPLIPASEC.
DR ProDom; PD001202; PI_PLC_Y; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00148; PLCX; 1.
DR SMART; SM00149; PLCYC; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR PROSITE; PS00007; PIPLC_X_DOMAIN; 1.
DR PROSITE; PS00008; PIPLC_Y_DOMAIN; 1.
SQ SEQUENCE 1128 AA; 125763 MW; 8598CC59D89ADCE5 CRC64;

Query Match 85.0%; Score 34; DB 11; Length 1128;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
DB 291 CHAVQC 296

RESULT 19
Q910CO PRELIMINARY; PRT; 130 AA.
ID Q910CO;
AC Q910CO;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE Hypothetical protein PA2722.
GN PA2722.

OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.

OX NCBI_TaxID=287;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004700; AAG06110.1; -
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 130 AA; 14641 MW; ECDEBCD7F247836E CRC64;

Query Match 82.5%; Score 33; DB 16; Length 130;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
|||:
DB 1 CHAVQC 6

```

RA Planagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens CS8.";
RL Science 294:2323-2328(2001).
DR EMBL; AB009899; AAL41173.1; ALT_INIT.
DR EMBL; AB007956; AAK85969.1; -.
KW Complete proteome.
SQ SEQUENCE 270 AA; 28869 MW; 58CB4BC556A1E279 CRC64;

Query Match      85.0%; Score 34; DB 16; Length 270;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 246 CHVDC 251

RESULT 14
O9GRX1 PRELIMINARY; PRT; 481 AA.
AC O9GRX1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GP11 protein.
GN GP11.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Campesán S., Chalmers D., Sandrelli F., Megighian A., Peixoto A.A.,
RA Costa R., Kyriacou C.P.;
RT "Comparative analysis of the nonA region in Drosophila identifies
RT highly diverged 5' gene that may constrain nonA promoter evolution.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ298995; CAC10056.1; -.
DR FlyBase; FBgn0042713; Gp11.
SQ SEQUENCE 481 AA; 55989 MW; 2E77ECA7643E8CB5 CRC64;

Query Match      85.0%; Score 34; DB 5; Length 481;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 445 CHEIDC 450

RESULT 15
O9VEU6 PRELIMINARY; PRT; 649 AA.
AC O9VEU6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG14885 protein.
GN CG14885.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

```

```

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
Deodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Swirskas R., Taft R., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AB003713; AAF55322.1; -.
DR HSP; P16068; 1AWN.
DR FlyBase; FBgn0038435; CG14885.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW Lyase.
SQ SEQUENCE 649 AA; 73892 MW; 3CF9F52B7E5D7505 CRC64;

Query Match      85.0%; Score 34; DB 5; Length 649;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
Db 26 CHIDC 31

RESULT 16
O8S315 PRELIMINARY; PRT; 745 AA.
AC O8S315;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative exo-1,3-beta-glucanase.
GN EXO1.
OS Phytophthora infestans (Potato late blight fungus).
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4787;
RN [1]
RP SEQUENCE FROM N.A.
RA McLeod A., Smart C.D., Fry W.E.;
RT "Characterization of 1,3-beta-glucanase and 1,3;1,4-beta-glucanase

```

DR PROSITE; PS01186; EGF_2; UNKNOWN 2.
 DR PROSITE; PS01248; LAMININ TYPE EGF; UNKNOWN 4.
 SQ SEQUENCE 1168 AA; 128899 MW; 66EB2B81AA232215 CRC64;

Query Match
 Best Local Similarity 87.5%; Score 35; DB 11; Length 1168;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
 |||||
 Db 425 CHACDC 430

RESULT 11
 ID Q9U3U7 PRELIMINARY; PRT; 1623 AA.
 AC Q9U3U7;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Laminin gamma 1 precursor.
 GN LANB2.
 OS Anopheles gambiae (African malaria mosquito).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
 CC Anopheles.
 OX NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=G3;
 RX MEDLINE=21126587; PubMed=11221310;
 RA Vlachou D., Lycett G., Siden-kiamos I., Blass C., Sindén R.E.,
 RA Louis C.;
 RT "Anopheles gambiae laminin interacts with the P25 surface protein of
 RT Plasmodium berghei ookinetes".
 RL Mol. Biochem. Parasitol. 112:229-237(2001).
 DR EMBL; AJ271193; CAB66001.1; -;
 DR HSSP; P02468; ITLE.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000034; Laminin B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001886; LamNT.
 DR Pfam; PF00052; laminin B; 1.
 DR Pfam; PF00053; laminin EGF; 10.
 DR Pfam; PF00055; laminin Nterm; 1.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR ProDom; PD002082; LamNT; 1.
 DR SMART; SM00180; EGF_Lam; 10.
 DR SMART; SM00281; LamB; 1.
 DR SMART; SM00136; LamNT; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN 8.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01248; LAMININ TYPE EGF; 11.
 KW EGF-like domain; Glycoprotein; Laminin EGF-like domain; Repeat;
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 1623 AA; 179542 MW; 7C87949E5867B702 CRC64;

Query Match
 Best Local Similarity 87.5%; Score 35; DB 5; Length 1623;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
 |||||
 Db 984 CHACDC 989

RESULT 12
 ID Q9FSL0 PRELIMINARY; PRT; 264 AA.
 AC Q9FSL0;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Thaumatin-like protein.
 GN H0806H05.110.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Han B., Peng Q., Huang Y.C., Chen Z.H., Zhou B., Li Y., Zhu J.J.,
 RA Tang Y.S., Zhao Q., Liu Y.L., Wu J., Yu Z., Fan D.L., Chen L.,
 RA Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Zhu J., Liu X.H., Hu X.,
 RA Lei H.Y., Zhang Y.J., Wang R., Li C., Lu Y., Chen X.C., Zhang Y.,
 RA Hu H., Jia P.X., Li T., Qian Y.M., Ying K., Hong G.F.;
 RT "Oryza sativa indica (Guangluai4) genomic DNA, chromosome 4, BAC
 RT clone: H0806H05".
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL442113; CAC09477.1; -;
 DR HSSP; P33679; IDU5.
 DR InterPro; IPR001938; Thaumatin.
 DR Pfam; PF00314; thaumatin; 1.
 DR PRINTS; PR00347; THAUMATIN.
 DR ProDom; PD001321; Thaumatin; 1.
 DR SMART; SM00205; THN; 1.
 DR PROSITE; PS00316; THAUMATIN; UNKNOWN 1.
 SQ SEQUENCE 264 AA; 26573 MW; 8C1D57B9F5BDD560 CRC64;

Query Match 85.0%; Score 34; DB 10; Length 264;
 Best Local Similarity 83.3%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
 |||||
 Db 167 CHAVSC 172

RESULT 13
 ID Q8UIZ2 PRELIMINARY; PRT; 270 AA.
 AC Q8UIZ2;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Enoyl-(acyl-carrier-protein) reductase.
 GN FABI OR ATU0149 OR AGR_C_242.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kuyavlin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58".
 RL Science 294:2317-2323 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,

FT TRANSMEM 444 461 BY SIMILARITY.
 FT TRANSMEM 485 504 BY SIMILARITY.
 FT TRANSMEM 625 646 BY SIMILARITY.
 FT TRANSMEM 1104 1123 BY SIMILARITY.
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 926 926 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1132 AA; 125791 MW; 9E1B12A58D796CF0 CRC64;

Query Match 90.0%; Score 36; DB 12; Length 1132;
 Best Local Similarity 83.3%; Pred. No. 44;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 Db 1043 CHGVDC 1048

RESULT 8

Q9Y7U5 PRELIMINARY; PRT; 1275 AA.
 ID Q9Y7U5
 AC Q9Y7U5
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Probable RHO1 GDP-GTP exchange protein 2.
 GN SPCC645.06C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Rajandream M.A., Barrell B.G., Rieger M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE EXCHANGE OF RHO1 GDP-BOUND FORM INTO
 GTP-BOUND FORM.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 DR EMBL; AL049498; CAB39902.1; -.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR00219; RhoGEF.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00621; RhoGEF; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RhoGEF; 1.
 DR PROSITE; PS00003; PH DOMAIN; 1.
 KW Guanine-nucleotide releasing factor.
 FT DOMAIN 692 855 PH.
 SQ SEQUENCE 1275 AA; 144430 MW; E5B6CFA7780FD2FF CRC64;

Query Match 90.0%; Score 36; DB 3; Length 1275;
 Best Local Similarity 66.7%; Pred. No. 49;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 Db 1099 CHSIDC 1104

RESULT 9

O45251 PRELIMINARY; PRT; 1404 AA.
 ID O45251
 AC O45251
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)

DE C14B4.2 protein.
 GN C14B4.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Basham V.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=94150718; PubMed=7906398;
 RX Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showlken R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans*;
 RL Nature 368:32-38(1994).
 DR EMBL; Z81471; CAB03887.1; -.
 DR EMBL; Z81483; CAB03887.1; JOINED.
 DR EMBL; Z81483; CAB03966.1; -.
 DR EMBL; Z81471; CAB03966.1; JOINED.
 SQ SEQUENCE 1404 AA; 155476 MW; BA379F704C02624F CRC64;

Query Match 90.0%; Score 36; DB 5; Length 1404;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 Db 843 CHVDC 848

RESULT 10

Q91V90 PRELIMINARY; PRT; 1168 AA.
 ID Q91V90
 AC Q91V90
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE BM106N23.2 (Laminin, beta 3) (Similar to laminin, beta 3).
 GN LAMB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kay M.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL365314; CAC41380.1; -.
 DR EMBL; BC008516; AA08516.1; -.
 DR MGI; MGI:99915; Lamb3.
 DR InterPro; IPR000345; CytC_heme_bind.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001886; LamNT.
 DR Pfam; PF00053; laminin_EGF; 6.
 DR Pfam; PF00055; laminin_Nterm; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_5.

Q9T2S1
ID Q9T2S1 PRELIMINARY; PRT; 589 AA.
AC Q9T2S1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Fibulin-1D (Fragment).
GN FBLN1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBI489 HIM-8;
RX MEDLINE=99120531; PubMed=9923656;
RA Barth J.L., Argraves K.M., Roark B.F., Little C.D., Argraves W.S.;
RT "Identification of chicken and C. elegans fibulin-1 homologs and
characterization of the C. elegans fibulin-1 gene.";
RL Matrix Biol. 17:635-646(1998).
DR EMBL; AF070477; AAC24035.1; -.
DR HSP; P16109; IFSB.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00179; EGF_CA; 5.
DR PROSITE; PS00010; ASX HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 7.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1
SQ SEQUENCE 589 AA; 63984 MW; 8EA3B8FCB0B97BE6 CRC64;

Query Match 90.0%; Score 36; DB 5; Length 589;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 446 CHSIDC 451

RESULT 6
Q18026 PRELIMINARY; PRT; 798 AA.
ID O18026
AC O18026; Q20903;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F56H11.1 protein.
GN F56H11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC Wilkinson J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Lloyd C.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68749; CA92962.1; -.
DR EMBL; Z68219; CA92962.1; JOINED.
DR EMBL; Z68219; CA92483.1; -.
DR EMBL; Z68749; CA92483.1; JOINED.
DR HSP; P16109; IFSB.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 2.
DR Pfam; PF00008; EGF; 5.
DR SMART; SM00104; ANATO; 2.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF-like; 6.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE; PS00010; ASX HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 8.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 798 AA; 87205 MW; 3BF1EB9ED54DBBF9 CRC64;

Query Match 90.0%; Score 36; DB 5; Length 798;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 632 CHSIDC 637

RESULT 7
Q07022 PRELIMINARY; PRT; 1132 AA.
ID Q07022
AC Q07022;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE M polyprotein precursor [CONTAINS: glycoprotein G1; glycoprotein G2].
OS Thailand virus.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Hantavirus.
OX NCBI_TaxID=28291;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=THAI749;
RX MEDLINE=94082451; PubMed=8259656;
RA Xiao S.Y., Leduc J.W., Chu Y.K., Schmaljohn C.S.;
RT "Phylogenetic analyses of virus isolates in the genus Hantavirus,
family Bunyaviridae.";
RL Virology 198:205-217(1994).
CC -1- MISCELLANEOUS: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE
CC -1- PROTEINS INCLUDING GLYCOPROTEIN G1 AND GLYCOPROTEIN G2.
CC -1- SIMILARITY: TO OTHER HANTAVIRUSES M SEGMENT PROTEINS.
DR EMBL; L08756; AAA16239.1; -.
DR InterPro; IPR002534; Hanta_G1.
DR InterPro; IPR002532; Hanta_G2.
DR InterPro; IPR000634; S/T dehydratase.
DR Pfam; PF01567; Hanta_G1; 1.
DR Pfam; PF01561; Hanta_G2; 1.
DR ProDom; PD001813; Hanta_G2; 1.
DR PROSITE; PS00165; DEHYDRATASE SER THR; 1.
KW Polyprotein; Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 16 BY SIMILARITY.
FT CHAIN 17 1132 M POLYPROTEIN.
FT CHAIN 17 646 GLYCOPROTEIN G1 (BY SIMILARITY).
FT CHAIN 647 1132 GLYCOPROTEIN G2 (BY SIMILARITY).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11133;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=QUEBEC;
 RA Yoo D., Pei Y., Parker M.D., Cox G.J.;
 RT "Bovine coronavirus (Quebec strain) full-length genomic sequence."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF220295; AAL40397.1;
 DR InterPro: IPR000606; Viral_helicase1.
 DR Pfam: PF01443; Viral_helicase1; 1.
 FT NON TER 1
 SQ SEQUENCE 2685 AA; 303641 MW; 6E0B02408605556 CRC64;

Query Match 92.5%; Score 37; DB 12; Length 2685;
 Best Local Similarity 83.3%; Pred. No. 61;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 243 CHALDC 248
 |||:|

RESULT 2

ID Q91A29 PRELIMINARY; PRT; 7094 AA.
 AC Q91A29;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Replicase.
 OS bovine coronavirus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCOV-ENT;
 RA Chouljenko V.N., Gorbalenya A.E., Lin X.Q., Storz J., Kousoulas K.G.;
 RT "Comparison of Genomic and Predicted Amino Acid Sequences of
 RT Respiratory and Enteric Bovine Coronaviruses Isolated From the Same
 RT Animal With Fatal Shipping Pneumonia."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE +
 CC NAD(+) = 3-PHOSPHO-D-GLYCEROYL PHOSPHATE + NADH.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR EMBL: AF391541; AAK83365.1; -;
 DR InterPro: IPR002589; Alpp.
 DR InterPro: IPR002086; Aldehyde dehydr.
 DR InterPro: IPR002705; Peptidase C16.
 DR InterPro: IPR000606; Viral_helicase1.
 DR Pfam: PF01661; Alpp; 1.
 DR Pfam: PF01831; Peptidase C16; 1.
 DR Pfam: PF01443; Viral_helicase1; 1.
 DR PROSITE: PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
 KW Glycolysis; NAD; Oxidoreductase.
 SQ SEQUENCE 7094 AA; 797292 MW; 5B02FC1059728FF3 CRC64;

Query Match 92.5%; Score 37; DB 12; Length 7094;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 4617 CHALDC 4622
 |||:|

RESULT 3

Q8V439 PRELIMINARY; PRT; 7094 AA.
 ID Q8V439
 AC Q8V439;

DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Replicase.
 OS bovine coronavirus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCOV-LUN;
 RA Chouljenko V.N., Gorbalenya A.E., Lin X.Q., Storz J., Kousoulas K.G.;
 RT "Comparison of Genomic and Predicted Amino Acid Sequences of
 RT Respiratory and Enteric Bovine Coronaviruses Isolated From the Same
 RT Animal With Fatal Shipping Pneumonia."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF391542; AAL57305.1; -;
 DR InterPro: IPR002589; Alpp.
 DR InterPro: IPR002086; Aldehyde dehydr.
 DR InterPro: IPR002705; Peptidase C16.
 DR InterPro: IPR000606; Viral_helicase1.
 DR Pfam: PF01661; Alpp; 1.
 DR Pfam: PF01831; Peptidase C16; 1.
 DR Pfam: PF01443; Viral_helicase1; 1.
 DR SMART: SM00506; Alpp; 1.
 DR PROSITE: PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
 SQ SEQUENCE 7094 AA; 797367 MW; 4FC564AE7C707E3 CRC64;

Query Match 92.5%; Score 37; DB 12; Length 7094;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 4617 CHALDC 4622
 |||:|

RESULT 4

ID Q9NWP6 PRELIMINARY; PRT; 299 AA.
 AC Q9NWP6;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE KIAA2204 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ILEAL MUCOSA;
 RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
 RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
 RA Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK000708; BAA91333.1; -;
 DR InterPro: IPR000050; PID_domain.
 DR SMART: SM00462; PTB; 1.
 DR PROSITE: PS01179; PID; 1.
 DR PROSITE: PS01179; PID; 1.
 SQ SEQUENCE 299 AA; 33816 MW; 02332293A2FD1F46 CRC64;

Query Match 90.0%; Score 36; DB 4; Length 299;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 246 CHAVEC 251
 |||:|

RESULT 5

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:42:49 ; Search time 31.6364 Seconds

(without alignments)
39.078 Million cell updates/sec

Title: US-10-105-008-20

Perfect score: 40

Sequence: 1 CHAVDC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 205047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SPTREMBL 21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	92.5	2685	12 Q8V6W6	Q8V6W6 bovine coro
2	37	92.5	7094	12 Q91A29	Q91A29 bovine coro
3	37	92.5	7094	12 Q8V439	Q8V439 bovine coro
4	36	90.0	299	4 Q9NWP6	Q9NWP6 homo sapien
5	36	90.0	589	5 Q9T2S1	Q9T2S1 caenorhabdi
6	36	90.0	798	5 O18026	O18026 caenorhabdi
7	36	90.0	1132	12 Q07022	Q07022 thailand vi
8	36	90.0	1275	3 Q9Y7U5	Q9Y7U5 schizosacch
9	36	90.0	1404	5 O45251	O45251 caenorhabdi
10	35	87.5	1168	11 Q91V90	Q91V90 mus musculu
11	35	87.5	1623	5 Q9U3U7	Q9U3U7 anopheles g
12	34	85.0	264	10 Q9FSL0	Q9FSL0 oryza sativ
13	34	85.0	270	16 Q8U122	Q8U122 agrobacteri
14	34	85.0	481	5 Q9GRX1	Q9GRX1 drosophila
15	34	85.0	649	5 Q9VEU6	Q9VEU6 drosophila
16	34	85.0	745	10 Q8S315	Q8S315 phytophthor

RESULT 1

Q8V6W6 PRELIMINARY; PRT; 2685 AA.
ID Q8V6W6;
AC Q8V6W6;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DB RNA polymerase lb (fragment)
OS Bovine coronavirus (strain Quebec).

ALIGNMENTS

17	34	85.0	1074	10	Q94H19	Q94H19 oryza sativ
18	34	85.0	1128	11	Q9QXG1	Q9QXG1 mus musculu
19	33	82.5	130	16	Q910C0	Q910C0 pseudomonas
20	33	82.5	138	16	Q98K76	Q98K76 rhizobium l
21	33	82.5	176	17	Q8TYK3	Q8TYK3 methanopyru
22	33	82.5	246	10	Q9ARG0	Q9ARG0 capsicum an
23	33	82.5	466	5	O96564	O96564 junonia coe
24	33	82.5	527	12	Q9E226	Q9E226 helicoverpa
25	33	82.5	715	5	Q9VDD2	Q9VDD2 drosophila
26	33	82.5	866	5	Q917E3	Q917E3 drosophila
27	33	82.5	933	5	Q917B4	Q917B4 drosophila
28	33	82.5	958	5	Q9V560	Q9V560 drosophila
29	33	82.5	1188	10	Q8SV59	Q8SV59 arabidopsis
30	32	80.0	134	16	Q98D49	Q98D49 rhizobium l
31	32	80.0	134	16	Q92SM1	Q92SM1 rhizobium m
32	32	80.0	152	16	Q92UX9	Q92UX9 rhizobium m
33	32	80.0	180	10	Q9FUN6	Q9FUN6 capsicum an
34	32	80.0	207	10	Q8S411	Q8S411 solanum nig
35	32	80.0	246	10	Q40529	Q40529 nicotiana t
36	32	80.0	248	10	Q9FT35	Q9FT35 solanum dul
37	32	80.0	302	10	Q9LN72	Q9LN72 arabidopsis
38	32	80.0	307	10	Q8VZ48	Q8VZ48 arabidopsis
39	32	80.0	314	3	Q96VC6	Q96VC6 zygosacchar
40	32	80.0	420	5	Q17468	Q17468 caenorhabdi
41	32	80.0	429	5	Q9V538	Q9V538 caenorhabdi
42	32	80.0	454	10	Q9SB62	Q9SB62 arabidopsis
43	32	80.0	610	6	Q95L61	Q95L61 equus cabal
44	31	77.5	38	16	Q9K2C4	Q9K2C4 chlamydia p
45	31	77.5	81	9	Q8W763	Q8W763 bacterioph
46	31	77.5	93	6	Q9GMK8	Q9GMK8 macaca fasc
47	31	77.5	111	10	Q8S0R8	Q8S0R8 oryza sativ
48	31	77.5	141	16	Q8XUD1	Q8XUD1 ralstonia s
49	31	77.5	160	2	Q87682	Q87682 arthrobacte
50	31	77.5	200	4	Q9BZHI	Q9BZHI homo sapien
51	31	77.5	200	4	Q9H3V2	Q9H3V2 homo sapien
52	31	77.5	223	16	Q984H8	Q984H8 rhizobium l
53	31	77.5	253	2	Q9KJ12	Q9KJ12 lactococcus
54	31	77.5	254	16	Q8ZEV5	Q8ZEV5 yersinia pe
55	31	77.5	263	10	Q9ARE0	Q9ARE0 flaveria tr
56	31	77.5	270	17	Q97YM2	Q97YM2 sulfolobus
57	31	77.5	286	2	Q9KIP8	Q9KIP8 escherichia
58	31	77.5	304	16	Q8Z011	Q8Z011 anabaena sp
59	31	77.5	391	16	Q8XHD7	Q8XHD7 clostridium
60	31	77.5	396	17	Q9HP92	Q9HP92 halobacteri
61	31	77.5	412	16	Q981T1	Q981T1 rhizobium l
62	31	77.5	416	17	Q30211	Q30211 archaeoglob
63	31	77.5	431	16	Q9K015	Q9K015 neisseria m
64	31	77.5	500	5	Q9NEG9	Q9NEG9 leishmania
65	31	77.5	508	10	Q9C804	Q9C804 arabidopsis
66	31	77.5	516	16	Q9HT19	Q9HT19 pseudomonas
67	31	77.5	601	16	Q9ZBJ9	Q9ZBJ9 streptomyce
68	31	77.5	646	13	Q919B9	Q919B9 torpedo mar
69	31	77.5	668	12	P89512	P89512 feline panl
70	31	77.5	668	12	P89513	P89513 feline panl
71	31	77.5	668	12	P89514	P89514 feline panl
72	31	77.5	668	12	P89515	P89515 feline panl
73	31	77.5	668	12	P89516	P89516 feline panl
74	31	77.5	668	12	Q84393	Q84393 canine parv
75	31	77.5	668	12	P90449	P90449 feline panl

Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7
|||||

Db 1224 HAVSSC 1229

RESULT 24

Q9WD02 ID Q9WD02 PRELIMINARY; PRT; 77 AA.
AC Q9WD02;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN NS3.
OS Hepatitis G virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC GBV-C/HGV group.
OX NCBI_TaxID=45255;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M13;
RX MEDLINE=99294779; PubMed=10364608;
RA Menendez C., Sanchez-Tapias J.M., Alonso P.L., Gimenez-Barcons M.,
RA Kahigwa E., Aponte J.J., Mshinda H., Navia M.M., Jimenez de Anta M.T.,
RA Rodes J., Saiz J.C.;
RT "Molecular evidence of mother-to-infant transmission of hepatitis G
RT virus among women without known risk factors for parenteral
RT infections.";
RL J. Clin. Microbiol. 37:2333-2336(1999).
DR EMBL; AF099460; AAD32059.1; -;
DR HSP; P27958; IAIU.
DR MEROPS; S29.002; -;
FT NON TER 1 1
FT NON TER 77 77
SQ SEQUENCE 77 AA; 8229 MW; B2F6B6AC21051B01 CRC64;

Query Match 76.2%; Score 32; DB 12; Length 77;

Best Local Similarity 57.1%; Pred. No. 38;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
|||:

Db 10 CHSVAC 16
|||:

RESULT 25

Q98434 ID Q98434 PRELIMINARY; PRT; 157 AA.
AC Q98434;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE A382R protein.
GN A382R.
OS Paramyxium bursaria chlorocella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96400190; PubMed=8806566;
RA Kutish G.F., Li Y., Lu Z., Furuta M., Rock D.L., Van Etten J.L.;
RT "Analysis of 76 kb of the chlorocella virus PBCV-1 330-kb genome: map
RT positions 182 to 258.";
RL Virology 223:303-317(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20013326; PubMed=10544099;
RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA Lisec A.D., Nickerson K.W., Van Etten J.L.;
RT "Chlorocella virus PBCV-1 encodes a functional homosperridine

synthase.";
RL Virology 263:254-262(1999).
RN [3]
RP SEQUENCE FROM N.A.

RX MEDLINE=20478054; PubMed=11021991;

RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;

RT "Characterization of a beta-1,3-glucanase encoded by chlorocella virus

RT PBCV-1.";

RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.

RA Van Etten J.L.;

RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

RN [5]
RP SEQUENCE FROM N.A.

RA Van Etten J.L.;

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [6]
RP SEQUENCE FROM N.A.

RA Van Etten J.L.;

RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

RN [7]
RP SEQUENCE FROM N.A.

RA Graves M.V., Van Etten J.L.;

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [8]
RP SEQUENCE FROM N.A.

RA Graves M.V., Van Etten J.L.;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [9]
RP SEQUENCE FROM N.A.

RA Gurnon J.R., Graves M.V., Van Etten J.L.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; U42580; AAC96750.1; -;
SQ SEQUENCE 157 AA; 18046 MW; 25EF8B6122E1014E CRC64;

Query Match 76.2%; Score 32; DB 12; Length 157;

Best Local Similarity 71.4%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
|||||

Db 102 CHAVRC 108
|||||

Search completed: June 20, 2003, 20:57:19

Job time : 41.9091 secs


```

ID Q855U4 PRELIMINARY; PRT; 728 AA.
AC Q855U4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical 81.3 kDa protein.
GN QJ123P12.6.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC104428; AAM15784.1; -
KW Hypothetical protein.
SQ SEQUENCE 728 AA; 81253 MW; 4B9462F820196914 CRC64;

Query Match 78.6%; Score 33; DB 10; Length 728;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 21 CHAFSEC 27

RESULT 22
Q9VRA4 PRELIMINARY; PRT; 1528 AA.
AC Q9VRA4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG1702 protein.
GN CG1702.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benton P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Feiraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
```

```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003571; AAF50899.1; -
DR HSP; P30712; ILJR.
DR FlyBase; FBgn0031117; CG1702.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR004045; GST_Nterm.
DR InterPro; IPR000195; RabGAP_TBC.
DR InterPro; IPR004012; Run.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
DR Pfam; PF02759; RUN; 1.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
SQ SEQUENCE 1528 AA; 171986 MW; F09084A35C9AF116 CRC64;

Query Match 78.6%; Score 33; DB 5; Length 1528;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 53 CHTVHSC 59

RESULT 23
Q95MJ0 PRELIMINARY; PRT; 2425 AA.
ID Q95MJ0;
AC Q95MJ0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Mannose-6-phosphate/insulin-like growth factor II receptor
DE (Fragment).
GN M6P/IGF2R.
OS Tupiaia glis (Tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupia.
OX NCBI_TaxID=9395;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313641; PubMed=11420613;
RA Killian J.K., Buckley T.R., Stewart N., Munday B.L., Jirtle R.L.;
RT "Marsupials and Eutherians reunited: genetic evidence for the Theria
RT hypothesis of mammalian evolution."
RL Mamm. Genome 12:513-517(2001).
DR EMBL; AF339161; AAK71867.1; -
DR InterPro; IPR000479; CIMR.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR001917; NHRtransf_2.
DR Pfam; PF00878; CIMR; 13.
DR Pfam; PF00040; fn2; 1.
DR PRODOM; PD000995; FN Type II; 1.
DR PROSITE; PS00599; AA TRANSFER CLASS 2; UNKNOWN 1.
DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 2425 2425
SQ SEQUENCE 2425 AA; 266339 MW; 40306E7E6380B386 CRC64;

Query Match 78.6%; Score 33; DB 6; Length 2425;
```

DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00608; LDLRA_2; 2.
KW Glycoprotein.

SQ SEQUENCE 570 AA; 64515 MW; 2AA95ED3CF38E942 CRC64;

Query Match 78.6%; Score 33; DB 5; Length 570;

Best Local Similarity 71.4%; Pred. No. 1.3e+02; Length 570;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

Db 422 CHWDSC 428

RESULT 18

Q9URJ8 PRELIMINARY; PRT; 605 AA.

AC Q9URJ8

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)

DE Glucose oxidase.

GN GOX

OS Penicillium amagasakiense.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.

OX NCBI_TaxID=63559;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33245;

RA Witt S., Kalisz H.M., Singh M.;

RT "Cloning, DNA sequencing and expression of the glucose oxidase gene

from Penicillium amagasakiense."

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF012277; AAD01493.1; -

DR HSSP; P81156; IGPE.

DR InterPro; IPR001017; GMC_oxred.

DR InterPro; IPR002016; Peroxidase.

DR Pfam; PF00732; GMC_oxred; 1.

DR PROSITE; PS00623; GMC_OXRED_1; 1.

DR PROSITE; PS00624; GMC_OXRED_2; 1.

DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN 1.

SQ SEQUENCE 605 AA; 65768 MW; F9F5E0A0687B8465 CRC64;

Query Match 78.6%; Score 33; DB 3; Length 605;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HAVSSC 7

Db 538 HAVSSC 543

RESULT 19

Q62285

AC Q62285 PRELIMINARY; PRT; 642 AA.

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

DE Uromodulin.

GN UMOD.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95178555; PubMed=7873609;

RA Prasad K., Bates J., Badgett A., Dell M., Sukhatme V., Yu H.,

RA Kumar S.;

RT "Nucleotide sequence and peptide motifs of mouse uromodulin (Tamm-

Horsfall protein)-the most abundant protein in mammalian urine.";

RL Biochim. Biophys. Acta 1260:328-332(1995).
DR EMBL; L33406; AAA73896.1; -
DR HSP; P35555; 1EMN.

DR MGD; MGI:102674; Umod.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR001507; Endoglin/CD105.

DR Pfam; PF00008; EGF; 3.

DR Pfam; PF00100; zona pellucida; 1.

DR PRINTS; PR00023; ZPELUUCIDA.

DR SMART; SM00179; EGF_Ca; 2.

DR SMART; SM00001; EGF-like; 1.

DR SMART; SM00241; ZP; 1.

DR PROSITE; PS00010; ASX_HYDROXYL; 2.

DR PROSITE; PS01186; EGF_2; 3.

DR PROSITE; PS01187; EGF_Ca; 2.

DR PROSITE; PS00682; ZP_DOMAIN; 1.

KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.

SQ SEQUENCE 642 AA; 70839 MW; 415A7E58D3D2B1A9 CRC64;

Query Match 78.6%; Score 33; DB 11; Length 642;

Best Local Similarity 57.1%; Pred. No. 1.4e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

Db 119 CHALATC 125

RESULT 20

Q91X17

AC Q91X17 PRELIMINARY; PRT; 642 AA.

DT 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

DE Similar to uromodulin.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RA Strausberg R.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC012973; AAL12973.1; -

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000345; CytC_heme_bind.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR001507; Endoglin/CD105.

DR Pfam; PF00008; EGF; 3.

DR Pfam; PF00100; zona pellucida; 1.

DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN 2.

DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.

DR PROSITE; PS01186; EGF_2; UNKNOWN 1.

DR PROSITE; PS01187; EGF_Ca; UNKNOWN 2.

DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN 1.

SQ SEQUENCE 642 AA; 70844 MW; 31B1461B4DCAE927 CRC64;

Query Match 78.6%; Score 33; DB 11; Length 642;

Best Local Similarity 57.1%; Pred. No. 1.4e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

Db 119 CHALATC 125

RESULT 21

Q8S504

Db 383 CHWVSC 389

Query Match 78.6%; Score 33; DB 5; Length 549;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
 DB 401 CHWVSC 407

RESULT 16

Q9VW30 ID Q9VW30 PRELIMINARY; PRT; 549 AA.
 AC Q9VW30
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE CG8756 protein.
 GN CG8756

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003516; AAF49122.1;
 DR FlyBase; FBgn0036901; CG8756.
 DR InterPro; IPR002557; Chitin_bind_Pera.
 DR InterPro; IPR002172; LDL_recept_A.
 DR Pfam; PF01607; CEM_14; 1.
 DR SMART; SM00494; ChtBD2; 1.
 DR SMART; SM00192; LDLA; 1.
 DR PROSITE; PS01209; LDLA_1; 1.
 DR PROSITE; PS00068; LDLA_2; 1.
 KW Glycoprotein.
 SQ SEQUENCE 549 AA; 62510 MW; 93C9B580BE255EE6 CRC64;

RESULT 17

Q9VW32 ID Q9VW32 PRELIMINARY; PRT; 570 AA.
 AC Q9VW32
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE CG8756 protein.
 GN CG8756

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003516; AAF49122.1;
 DR FlyBase; FBgn0036901; CG8756.
 DR InterPro; IPR002557; Chitin_bind_Pera.
 DR InterPro; IPR002172; LDL_recept_A.
 DR Pfam; PF01607; CEM_14; 2.
 DR SMART; SM00494; ChtBD2; 1.
 DR SMART; SM00192; LDLA; 1.

RESULT 13

Q8T2K0 PRELIMINARY; PRT; 495 AA.
 AC Q8T2K0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE Hypothetical 57.8 kDa protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Gloeckner G., Richinger L., Szafranski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.P., Guigo R., Kumpf K.,
 RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC115599; AAL92330.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 495 AA; 57844 MW; 21EBE7655DB89190 CRC64;

Query Match 78.6%; Score 33; DB 5; Length 495;
 Best Local Similarity 57.1%; Pred. No. 1.1e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

Db 346 CHEISNC 352

RESULT 14

Q9ERP7 PRELIMINARY; PRT; 527 AA.
 AC Q9ERP7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Intracellular adhesion molecule 1.
 GN ICAM1.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gaglia J.L., Greenfield E.A., Matteo A., Sharpe A.H., Freeman G.J.,
 RA Kuchroo V.K.;
 RT "ICAM-1 is critical for activation of CD28-deficient T cells";
 RL J. Immunol. 0:0-0(2000).
 DR EMBL; AF308456; AAG30280.1; -;
 DR HSP; P05362; 11C1.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003600; IG like.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00410; IG like; 1.
 SQ SEQUENCE 527 AA; 57598 MW; EA5385BF4D993050 CRC64;

Query Match 78.6%; Score 33; DB 11; Length 527;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CHAVSS 6

Db 454 CHAVSS 459

RESULT 15

Q9VW31 PRELIMINARY; PRT; 531 AA.
 AC Q9VW31;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE C88756 protein.
 GN C88756.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Posler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003516; AAF49123.2; -;
 DR FlyBase; FBgn0036901; CG8756.
 DR InterPro; IPR002557; Chitin_bind_Pera.
 DR InterPro; IPR002172; LDL_recept_A.
 DR Pfam; PF01607; CBM 14; 1.
 DR Pfam; PF00057; ldl_recept; 1.
 DR SMART; SM00494; ChtBD2; 1.
 DR SMART; SM00192; LDLA; 1.
 DR PROSITE; PS01209; LDLRA_1; 1.
 DR PROSITE; PS00688; LDLRA_2; 1.
 KW Glycoprotein.
 SQ SEQUENCE 531 AA; 60200 MW; 44301CA786A1A748 CRC64;

Query Match 78.6%; Score 33; DB 5; Length 531;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7

```

ID Q9XTJ7 PRELIMINARY; PRT; 709 AA.
AC Q9XTJ7;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Variant-specific surface protein (Fragment).
GN VSP417-6.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIS-136;
RX MEDLINE=99026095; PubMed=9806870;
RA EY P.L., Darby J.M.;
RT "Giardia intestinalis: conservation of the variant-specific surface protein VSP417-1 (TSA417) and identification of a divergent homologue encoded at a duplicated locus in genetic group II isolates.";
RL Exp. Parasitol. 90:250-261(1998).
DR EMBL; U89266; AAD03483.1; -.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR005127; Giardia_VSP.
DR InterPro; IPR001239; Kazal_inhib.
DR Pfam; PF03302; VSP; 2.
DR PRINTS; PRO0290; KAZALINHTR.
DR SMART; SM00261; FU; 3.
DR NON TER 709 709
FT SEQUENCE 709 AA; 71516 MW; 3512BB844B38D134 CRC64;

Query Match 81.0%; Score 34; DB 5; Length 709;
Best Local Similarity 85.7%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 327 CEAVSSC 333

RESULT 10
Q8X2P6 PRELIMINARY; PRT; 107 AA.
AC Q8X2P6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein EC83179.
GN EC83179.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuwara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; AP002561; BAB36602.1; -.
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 11984 MW; C74F73BE5D297112 CRC64;

Query Match 78.6%; Score 33; DB 16; Length 107;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 95 CHAITPC 101

```

```

RESULT 11
Q910I3 PRELIMINARY; PRT; 170 AA.
AC Q910I3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TRL10.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TS40E, AND 36YUEK;
RA Spaderna S., Blessing H., Bogner E., Britt W., Mach M.;
RT "Identification of glycoprotein gpTRL10 as a structural component of human cytomegalovirus.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF432092; AAL27474.1; -.
DR EMBL; AF432086; AAL27468.1; -.
SQ SEQUENCE 170 AA; 18940 MW; B496EE5601E33739 CRC64;

Query Match 78.6%; Score 33; DB 12; Length 170;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 59 CHAVCKC 65

```

```

RESULT 12
Q9BHU4 PRELIMINARY; PRT; 490 AA.
AC Q9BHU4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical predicted transmembrane protein P1046.57, unknown
DE function.
GN P1046.57.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Aert R., Robben J., Volckaert G., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL359683; CAC37162.2; -.
KW Transmembrane.
SQ SEQUENCE 490 AA; 54220 MW; D4B1922EBDA2795 CRC64;

Query Match 78.6%; Score 33; DB 5; Length 490;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CHAVSSC 7
Db 139 CHLTSSC 145

```

Db 427 CHAVGAC 433

RESULT 5

Q25668 Q25668 PRELIMINARY; PRT; 1766 AA.
 AC Q25668;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE Merozoite surface protein-1 precursor.
 OS Plasmodium chabaudi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5825;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AS;
 RX MEDLINE=94187794; PubMed=7511214;
 RA McKean P.G., O'Dea K., Brown K.N.;
 RT "Nucleotide sequence analysis and epitope mapping of the merozoite surface protein 1 from Plasmodium chabaudi AS.";
 RL Mol. Biochem. Parasitol. 62:199-210(1993).
 DR EMBL; L22982; AAA29499.1;
 KW Signal.
 FT SIGNAL.
 FT CHAIN 1 19 POTENTIAL.
 FT CHAIN 20 1766 MEROZOITE SURFACE PROTEIN-1.
 SQ SEQUENCE 1766 AA; 197557 MW; 6CFE98E26EDBD4C CRC64;

Query Match 83.3%; Score 35; DB 5; Length 1766;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
 ||: |||
 Db 402 CHSTSSC 408

RESULT 6

Q25685 Q25685 PRELIMINARY; PRT; 1785 AA.
 AC Q25685;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
 DE Major merozoite surface antigen precursor.
 OS Plasmodium chabaudi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5825;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IPP-C1;
 RX MEDLINE=91218805; PubMed=2090945;
 RA Deleersnijder W., Hendrix D., Bendahman N., Hanegeefs J., Brijs L.,
 RA Hamers-Casterman C., Hamers R.;
 RT "Molecular cloning and sequence analysis of the gene encoding the major merozoite surface antigen of Plasmodium chabaudi IP-PC1.";
 RT Mol. Biochem. Parasitol. 43:231-244(1990).
 RL EMBL; M34947; AAA29730.1;
 DR Merozoite; Signal.
 KW Merozoite; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1785 MAJOR MEROZOITE SURFACE ANTIGEN.
 SQ SEQUENCE 1785 AA; 198887 MW; AC5307F90113F4F7 CRC64;

Query Match 83.3%; Score 35; DB 5; Length 1785;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
 ||: |||
 Db 422 CHSTSSC 428

RESULT 7

Q9ARE0 Q9ARE0 PRELIMINARY; PRT; 263 AA.
 AC Q9ARE0;
 DT 01-JUN-2001 (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
 DE 2F-HD homeobox protein.
 GN HBI.
 OS Plavaria trineria (Clustered yellowtops).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 OC Heliaceae; Flaveria.
 OX NCBI_TaxID=4227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RX MEDLINE=21183146; PubMed=11289511;
 RA Windhoevel A., Hein I., Dabrowa R., Stockhaus J.;
 RT "Characterization of a novel class of plant homeodomain proteins that bind to the C4 phosphoenolpyruvate carboxylase gene of Flaveria trineria.";
 RT Plant Mol. Biol. 45:201-214(2001).
 RL EMBL; Y18577; CAC34413.1;
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 263 AA; 28903 MW; 4EBCCBD6D3FC131B CRC64;

Query Match 81.0%; Score 34; DB 10; Length 263;
 Best Local Similarity 71.4%; Pred. No. 43;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
 ||: |||
 Db 55 CHAVDGC 61

RESULT 8

Q9PY94 Q9PY94 PRELIMINARY; PRT; 394 AA.
 AC Q9PY94;
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE Nonstructural protein NSP1.
 GN NSP1.
 OS Human rotavirus (group C / strain Bristol).
 OC Viruses; GBRNA viruses; Reoviridae; Rotavirus.
 OX NCBI_TaxID=31567;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL;
 RX MEDLINE=20036633; PubMed=10567650;
 RA James V.L.A., Lambden P.R., Deng Y., Caul E.O., Clarke I.N.;
 RT "Molecular characterisation of human group C rotavirus genes 6, 7 and 9.";
 RT J. Gen. Virol. 80:3181-3187(1999).
 RL EMBL; AJ132204; CAB52752.1;
 DR InterPro; IPR002148; RotaNS53.
 DR ProDom; PD001671; Rota_NS53; 1.
 SQ SEQUENCE 394 AA; 46642 MW; 4FAB1BB841EDCF29 CRC64;

Query Match 81.0%; Score 34; DB 12; Length 394;
 Best Local Similarity 71.4%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CHAVSSC 7
 ||: |||
 Db 274 CHVWSEC 280

RESULT 9

Q9XTJ7

OS Polysphondylium pallidum.
 OC Eukaryota; Mycetozoa; Dictyostelida; Polysphondylium.
 OX NCBI_TaxID=13642;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=WS320;
 RX MEDLINE=94103266; PubMed=8276946;
 RA Manabe R., Saito T., Kumazaki T., Sakaitani S., Nakata N., Ochiai H.;
 RT "Molecular cloning and the COOH-terminal processing of gp64, a
 RT putative cell-cell adhesion protein of the cellular slime mold
 RT Polysphondylium pallidum.";
 RL J. Biol. Chem. 269:528-535(1994).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=WS320;
 RX MEDLINE=20011280; PubMed=10542319;
 RA Takaoka N., Fukuzawa M., Saito T., Sakaitani T., Ochiai H.;
 RT "Promoter analysis of the membrane protein gp64 gene of the cellular
 RT slime mold Polysphondylium pallidum.";
 RL Biochim. Biophys. Acta 1447:226-230(1999).
 DR EMBL; D14993; BAA03637.1; -;
 DR EMBL; AB027502; BAA06631.1; -;
 KW GPI-anchor; Signal.
 FT CHAIN 1 19
 FT SIGNAL 20 299 GP64.

FT CHAIN 320 AA; 32724 MW; APE158C56845EFAD CRC64;
 SQ SEQUENCE 320 AA; 32724 MW; APE158C56845EFAD CRC64;

Query Match 88.1%; Score 37; DB 5; Length 320;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CHAVSSC 7
 |||:|
 Db 165 CHPVSSC 171

RESULT 2

O91TP2 PRELIMINARY; PRT; 1055 AA.
 AC Q91TP2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE T47.
 OS Tupaia herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae.
 OX NCBI_TaxID=10397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2;
 RX MEDLINE=21211637; PubMed=11312357;
 RA Bahr U., Darai G.;
 RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree
 RT Shrew) Herpesvirus.";
 RL J. Virol. 75:4854-4870(2001).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=2;
 RA Darai G., Bahr U.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF291817; AAK57095.1; -;
 SQ SEQUENCE 1055 AA; 117780 MW; CCAS1E16FD38C39 CRC64;

Query Match 88.1%; Score 36; DB 12; Length 1055;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CHAVSSC 7
 |||:|
 Db 778 CHAVAAC 784

RESULT 3

O8TX27 PRELIMINARY; PRT; 444 AA.
 AC Q8TX27;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Uncharacterized membrane protein specific for M.kandleri, MK-15
 DE family.
 GN MK0851.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malykh A.G., Koonin E.V., Kozlyavkin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010375; AA02064.1; -;
 KW Complete proteome.

DR EMBL; AE010375; AA02064.1; -;
 KW Complete proteome.

SQ SEQUENCE 444 AA; 48310 MW; DCFE5233AAEC3711 CRC64;

Query Match 83.3%; Score 35; DB 17; Length 444;
 Best Local Similarity 71.4%; Pred. No. 43;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CHAVSSC 7
 |||:|
 Db 113 CHAVKAC 119

RESULT 4

O9KBP4 PRELIMINARY; PRT; 508 AA.
 AC Q9KBP4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein BH1881.
 GN BH1881.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001513; BAB05600.1; -;
 DR Interpro; IPR002821; Hydantoinease_A.
 DR Pfam; PF01968; Hydantoinease_A; 1.
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 508 AA; 56313 MW; 34E7BD1B5E44CE49 CRC64;

Query Match 83.3%; Score 35; DB 16; Length 508;
 Best Local Similarity 71.4%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CHAVSSC 7
 |||:|

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 20:42:49 ; Search time 36.9091 Seconds
(without alignments)
39.078 Million cell updates/sec

Title: US-10-105-008-44

Perfect score: 42

Sequence: 1 CHAVSSC 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

SPREMBL 21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	88.1	320	5 Q52085	Q52085 polysphondy
2	36	85.7	1055	12 Q91TP2	Q91TP2 tupaia herp
3	35	83.3	444	17 Q8TX27	Q8TX27 methanopyru
4	35	83.3	508	16 Q9KBP4	Q9KBP4 bacillus ha
5	35	83.3	1766	5 Q25668	Q25668 plasmodium
6	35	83.3	1785	5 Q25685	Q25685 plasmodium
7	34	81.0	263	10 Q9ARE0	Q9ARE0 flaveria tr
8	34	81.0	394	12 Q9PY94	Q9PY94 human rotav
9	34	81.0	709	5 Q9XJ77	Q9XJ77 giardia lam
10	33	78.6	107	16 Q8X2P6	Q8X2P6 escherichia
11	33	78.6	170	12 Q910I3	Q910I3 human cytom
12	33	78.6	490	5 Q9BHU4	Q9BHU4 leishmania
13	33	78.6	495	5 Q8T2K0	Q8T2K0 dictyosteli
14	33	78.6	527	11 Q9ERF7	Q9ERF7 cricetus
15	33	78.6	531	5 Q9VW31	Q9VW31 drosophila
16	33	78.6	549	5 Q9VW30	Q9VW30 drosophila

RESULT 1

ID	Q52085	PRELIMINARY;	PRT;	320 AA.
AC	Q52085; Q9U8R5;			Created
DT	01-NOV-1996 (T-EMBLrel. 01,			Last sequence update)
DT	01-NOV-1996 (T-EMBLrel. 01,			Last sequence update)
DT	01-MAR-2002 (T-EMBLrel. 20,			Last annotation update)
DE	GP64 precursor.			
GN	C-P644.			

ALIGNMENTS

17	33	78.6	570	5	Q9VW32	Q9VW32 drosophila
18	33	78.6	605	3	Q9URJ8	Q9URJ8 penicillium
19	33	78.6	642	11	Q62285	Q62285 mus musculus
20	33	78.6	642	11	Q91X17	Q91X17 mus musculus
21	33	78.6	728	10	Q8SSU4	Q8SSU4 oryza sativ
22	33	78.6	728	5	Q9VRA4	Q9VRA4 drosophila
23	33	78.6	2425	6	Q9SMJ0	Q9SMJ0 tupaia glis
24	32	76.2	77	12	Q9WD02	Q9WD02 hepatitis g
25	32	76.2	157	12	Q98434	Q98434 paramecium
26	32	76.2	251	12	Q9QH64	Q9QH64 gallid herp
27	32	76.2	261	16	Q9CP79	Q9CP79 pasteurella
28	32	76.2	349	5	Q9USN7	Q9USN7 leishmania
29	32	76.2	815	5	Q8T686	Q8T686 dictyosteli
30	32	76.2	976	10	Q9LUJ7	Q9LUJ7 arabidopsis
31	32	76.2	983	10	Q04327	Q04327 arabidopsis
32	32	76.2	1460	4	Q9COA1	Q9COA1 homo sapien
33	32	76.2	1952	5	Q95SN5	Q95SN5 drosophila
34	32	76.2	2146	3	Q95997	Q95997 aspergillus
35	32	76.2	2146	3	Q60026	Q60026 aspergillus
36	32	76.2	4547	5	Q9W343	Q9W343 drosophila
37	31.5	75.0	702	10	Q9LNL6	Q9LNL6 arabidopsis
38	31	73.8	88	9	Q8SC54	Q8SC54 stx2 conver
39	31	73.8	159	11	Q60691	Q60691 mus musculus
40	31	73.8	231	11	Q9CS11	Q9CS11 mus musculus
41	31	73.8	238	16	Q8UKG9	Q8UKG9 agrobacteri
42	31	73.8	256	5	Q8T2J8	Q8T2J8 dictyosteli
43	31	73.8	272	11	Q9CRT4	Q9CRT4 mus musculus
44	31	73.8	287	17	Q8TJ77	Q8TJ77 pyrobaculum
45	31	73.8	288	13	Q9DGH2	Q9DGH2 brachydanio
46	31	73.8	327	13	Q9DGH1	Q9DGH1 brachydanio
47	31	73.8	366	10	Q9LKG6	Q9LKG6 arabidopsis
48	31	73.8	382	13	Q9W7G2	Q9W7G2 gallus gall
49	31	73.8	391	11	Q9ER88	Q9ER88 mus musculus
50	31	73.8	396	11	Q9D5V9	Q9D5V9 mus musculus
51	31	73.8	396	11	Q8VCK8	Q8VCK8 mus musculus
52	31	73.8	494	5	Q9VWU0	Q9VWU0 drosophila
53	31	73.8	505	12	Q89768	Q89768 african swi
54	31	73.8	548	5	Q27169	Q27169 paramecium
55	31	73.8	549	13	Q9PVN5	Q9PVN5 xenopus lae
56	31	73.8	621	10	Q49047	Q49047 nicotiana t
57	31	73.8	704	5	Q9U048	Q9U048 giardia lam
58	31	73.8	709	5	Q97444	Q97444 giardia lam
59	31	73.8	815	5	Q44385	Q44385 caenorhabdi
60	31	73.8	826	13	Q9PVN6	Q9PVN6 xenopus lae
61	31	73.8	865	5	Q9N998	Q9N998 leishmania
62	31	73.8	898	13	Q12958	Q12958 oryzias lat
63	31	73.8	1001	10	Q9AX83	Q9AX83 oryza sativ
64	31	73.8	1053	4	Q9U0Q4	Q9U0Q4 homo sapien
65	31	73.8	1061	13	Q9PVN7	Q9PVN7 xenopus lae
66	31	73.8	1081	5	Q9U631	Q9U631 drosophila
67	31	73.8	1083	5	Q9VTT0	Q9VTT0 drosophila
68	31	73.8	1088	17	Q28333	Q28333 archaeglob
69	31	73.8	1139	5	Q9U280	Q9U280 caenorhabdi
70	31	73.8	1261	13	Q90ZM5	Q90ZM5 gallus gall
71	31	73.8	1272	13	Q9EPW7	Q9EPW7 mus musculus
72	31	73.8	1308	13	Q9DF77	Q9DF77 gallus gall
73	31	73.8	1350	13	Q91929	Q91929 xenopus lae
74	31	73.8	1377	13	Q9DDN5	Q9DDN5 xenopus lae
75	31	73.8	1736	4	Q8TES1	Q8TES1 homo sapien

DR PROSITE; PS00118; PA2_HIS; FALSE_NEG.
 DR PROSITE; PS00119; PA2_ASP; 1;

KW Glycoprotein. 26 120 BY SIMILARITY.
 FT DISULFID 28 44 BY SIMILARITY.
 FT DISULFID 43 99 BY SIMILARITY.
 FT DISULFID 49 127 BY SIMILARITY.
 FT DISULFID 50 92 BY SIMILARITY.
 FT DISULFID 59 85 BY SIMILARITY.
 FT DISULFID 78 90 BY SIMILARITY.
 FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 127 AA; 14630 MW; 996A448766859BCE CRC64;

Query Match 77.5%; Score 31; DB 1; Length 127;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CHAVDC 6
 ||: ||
 Db 44 CHSQDC 49

Search completed: June 20, 2003, 20:53:04
 Job time : 9.22727 secs

NCBI_TaxID=10797;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=90010964; PubMed=2794971;
RX Ran A.I., Mancius J.J., Diaz-Aroca E., Casal J.I.;
RT "Porcine parvovirus: DNA sequence and genome organization.";
RL J. Gen. Virol. 70:2541-2553(1989).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=91021005; PubMed=2219713;
RX Vasudevaacharya J., Basak S., Srinivas R.V., Compans R.W.;
RT "The complete nucleotide sequence of an infectious clone of porcine
parvovirus, strain NADL-2";
RL Virology 178:611-616(1990).
[3]
SEQUENCE OF 367-660 FROM N.A.
RX MEDLINE=90085785; PubMed=2596019;
RX Vasudevaacharya J., Basak S., Srinivas R.V., Compans R.W.;
RT "Nucleotide sequence analysis of the capsid genes and the right-hand
terminal palindromic of porcine parvovirus, strain NADL-2";
RL Virology 173:368-377(1989).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES NONCAPSID PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D00623; BAA00501.1; -;
CC EMBL; M38367; AAA46920.1; -;
CC EMBL; M32787; AAA46916.1; -;
CC PIR; A33302; UYVPP.
CC PIR; A36217; UYVNA.
CC InterPro; IPR001257; Parvo_NSI.
CC Pfam; PF01057; Parvo_NSI; 1.
CC Nonstructural protein; Noncapsid protein; DNA replication;
ATP-binding.
FT NP BIND 398 405 ATP (POTENTIAL).
FT CONFLICT 86 86 G -> R (IN REF. 2).
FT CONFLICT 274 274 K -> R (IN REF. 2).
FT CONFLICT 376 376 C -> V (IN REF. 2).
FT CONFLICT 621 634 TALTOHARFSTNIDT -> NLHLTPPDSAIRTP (IN
REF. 2).
SQ SEQUENCE 660 AA; 75300 MW; COB1DF2226A2EF0A CRC64;

Query Match 80.0%; Score 32; DB 1; Length 660;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
|||:|
Db 377 CHAIRC 382

RESULT 24
VNC5_PAVPK STANDARD; PRT; 662 AA.
AC PS2502;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN NS1.
OS Porcine parvovirus (strain Kresse) (PPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=73487;
[1]
PP SEQUENCE FROM N.A.
RX MEDLINE=96183900; PubMed=8642680;

Bergeron J., Hebert B., Tijssen P.;
RT "Genome organization of the Kresse strain of porcine parvovirus:
RT identification of the allotropic determinant and comparison with
RT those of NADL-2 and field isolates";
RL J. Virol. 70:2508-2515(1996).
CC -!- FUNCTION: SEEMS NECESSARY FOR VIRAL DNA REPLICATION.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES NONCAPSID PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U44978; AAC40229.1; -;
CC InterPro; IPR001257; Parvo_NSI.
CC Pfam; PF01057; Parvo_NSI; 1.
CC Nonstructural protein; Noncapsid protein; DNA replication;
ATP-binding.
FT NP BIND 398 405 ATP (POTENTIAL).
SQ SEQUENCE 662 AA; 75591 MW; B53F76D9F9FD613 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 662;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
|||:|
Db 377 CHAIRC 382

RESULT 25
PA2H_XENLA STANDARD; PRT; 127 AA.
AC P41485;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phospholipase A2 homolog otoconin-22 (Oc22).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
RN SEQUENCE.
RX MEDLINE=93264410; PubMed=8494877;
RA Pote K.G., Hauer C.R. III, Michel H., Shabanowitz J., Hunt D.P.,
RA Kretsinger R.H.;
RT "Otoconin-22, the major protein of aragonitic frog otoconia, is a
RT homolog of phospholipase A2";
RL Biochemistry 32:5017-5024(1993).
CC -!- FUNCTION: MAJOR PROTEIN OF THE ARAGONITIC OTOCONIA. IT IS UNLIKELY
CC THAT THIS PROTEIN HAS PHOSPHOLIPASE A2 ACTIVITY.
CC -!- SUBUNIT: MONOMER (PROBABLE).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: OTOCONIAL MEMBRANE IN THE MACULAE OF THE
CC SACCULE AND UTRICLE. OTOCONIA ARE COMPOSITES OF PROTEINS AND
CC INORGANIC CRYSTALS FORMED IN THE PERIPHERAL PORTION OF THE
CC VESTIBULAR SYSTEM OF VERTEBRATES. THE OTOCONIAL MEMBRANES CONTAIN
CC SMALL CRYSTALS OF CALCIUM CARBONATE KNOWN AS OTOLITHS (EAR STONES)
CC IF THERE IS A SINGLE DEPOSIT OR AS OTOCONIA (EAR DUST) IF THERE
CC ARE MANY. EACH MINERAL POLYMORPH OF OTOCONIA HAS A PROTEIN UNIQUE
CC TO THAT POLYMORPH.
CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
CC PIR; A49269; A49269.
DR HSSP; P00593; 4BP2.
DR InterPro; IPR001211; PhospholipaseA2.
DR Pfam; PF00068; phoslip; 1.
DR ProDom; PD000303; PhospholipaseA2; 1.
DR SMART; SMO00085; PA2c; 1.

FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 549 AA; 60079 MW; 85CEECDB7B0144C8 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 549;
 Best Local Similarity 66.7%; Pred. No. 66;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 ||| |:
 Db 238 CHVVEC 243

RESULT 22
 LEM2 MOUSE STANDARD; PRT; 612 AA.
 AC Q00690;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
 DE (ELAM-1) (leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
 DE (CD62E).
 GN SELE OR ELAM-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92283265; PubMed=1375914;
 RA Belkner-Andre M., van Huijsduijn R.H., Losberger C., Whelan J.,
 RA Delamarter J.P.;
 RT "Murine endothelial leukocyte-adhesion molecule 1 is a close
 structural and functional homologue of the human protein.";
 RL Eur. J. Biochem. 206:401-411(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92340571; PubMed=1378846;
 RA Weller A., Isenmann S., Vestweber D.;
 RT "Cloning of the mouse endothelial selectins. Expression of both E-
 and P-selectin is inducible by tumor necrosis factor alpha.";
 RL J. Biol. Chem. 267:15176-15183(1992).
 CC -!- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
 POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 GLYCOLIPIDS).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE SELECTIN/LECAM FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M80778; AAA37547.1; -;
 DR EMBL; M87862; AAA37577.1; ALT_INIT.
 DR HSP; P16581; 1KJA.
 DR MGD; MGI:98278; Sele.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001304; Lectin C.
 DR InterPro; IPR002396; Selectin.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00008; EGF; 1.

DR Pfam; PF00059; lectin_c; 1.
 DR Pfam; PF00084; sushi; 6.
 DR PRINTS; PRO0343; SELECTIN.
 DR SMART; SM00032; CCP; 6.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS06615; C-TYPE LECTIN_1; 1.
 DR PROSITE; PS06044; C-TYPE LECTIN_2; 1.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Signal; Sushi; Repeat.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT CHAIN 22 612 E-SELECTIN.
 FT DOMAIN 22 557 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 558 579 POTENTIAL.
 FT DOMAIN 580 612 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 38 138 C-TYPE LECTIN (SHORT FORM).
 FT DOMAIN 139 175 EGF-LIKE.
 FT DOMAIN 179 239 SUSHI 1.
 FT DOMAIN 242 301 SUSHI 2.
 FT DOMAIN 304 364 SUSHI 3.
 FT DOMAIN 367 427 SUSHI 4.
 FT DOMAIN 430 490 SUSHI 5.
 FT DOMAIN 493 549 SUSHI 6.
 FT DISULFID 40 138 BY SIMILARITY.
 FT DISULFID 111 130 BY SIMILARITY.
 FT DISULFID 143 154 BY SIMILARITY.
 FT DISULFID 148 163 BY SIMILARITY.
 FT DISULFID 165 174 BY SIMILARITY.
 FT DISULFID 180 225 BY SIMILARITY.
 FT DISULFID 210 238 BY SIMILARITY.
 FT DISULFID 243 287 BY SIMILARITY.
 FT DISULFID 273 300 BY SIMILARITY.
 FT DISULFID 305 350 BY SIMILARITY.
 FT DISULFID 336 363 BY SIMILARITY.
 FT DISULFID 368 413 BY SIMILARITY.
 FT DISULFID 399 426 BY SIMILARITY.
 FT DISULFID 431 476 BY SIMILARITY.
 FT DISULFID 462 489 BY SIMILARITY.
 FT DISULFID 494 535 BY SIMILARITY.
 FT DISULFID 521 548 BY SIMILARITY.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 612 AA; 66749 MW; 86F05713F0EC3C3D CRC64;

Query Match 80.0%; Score 32; DB 1; Length 612;
 Best Local Similarity 66.7%; Pred. No. 72;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 ||| |:
 Db 238 CHVVEC 243

RESULT 23
 VNCS PAVPN STANDARD; PRT; 660 AA.
 ID VNCS PAVPN STANDARD; PRT; 660 AA.
 AC P18547; P22965;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Noncapsid protein NS-1 (Nonstructural protein NS1).
 GN NS1.
 OS Porcine parvovirus (strain NADL-2) (PPV).
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

TX MEDLINE=95357444; PubMed=7630973;
RA Zhu B., Chen T.H.H., Li P.H.;
RT "Activation of two osmotin-like protein genes by abiotic stimuli and
RN fungal pathogens in transgenic potato plants.";
RL Plant Physiol. 108:929-937(1995).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95306785; PubMed=7787181;
RA Zhu B., Chen T.H.H., Li P.H.;
RT "Expression of three osmotin-like protein genes in response to
RN osmotic stress and fungal infection in potato.";
RL Plant Mol. Biol. 28:17-26(1995).
CC -1- INDUCTION: BY ASCISIC ACID (ABA), SALT, SALICYLIC ACID, WOUNDING,
CC AND FUNGAL INFECTION.
CC -1- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X72927; CRA51431.1; -
CC EMBL; X72926; CRA51430.1; -
CC HSP; P25871; IAUO.
CC InterPro; IPR001938; Thaumatin.
CC Pfam; PF00314; thaumatin; 1.
CC PRINTS; PR00347; THAUMATIN.
CC ProDom; PD001321; Thaumatin; 1.
CC SMART; SM00205; THN; 1.
CC PROSITE; PS00316; THAUMATIN; 1.
CC Signal.
CC SIGNAL 1 21
CC CHAIN 22 247
CC DISULFID 30 225
CC BY SIMILARITY.
CC DISULFID 72 82
CC BY SIMILARITY.
CC DISULFID 87 93
CC BY SIMILARITY.
CC DISULFID 141 213
CC BY SIMILARITY.
CC DISULFID 146 196
CC BY SIMILARITY.
CC DISULFID 154 164
CC BY SIMILARITY.
CC DISULFID 168 177
CC BY SIMILARITY.
CC DISULFID 178 183
CC BY SIMILARITY.
CC CONFLICT 92 92
CC Q -> L (IN REF. 2).
SQ SEQUENCE 247 AA; 26659 MW; BBCD9298CD81FBA1 CRC64;
Query Match 80.0%; Score 32; DB 1; Length 247;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CHAVDC 6
Db 141 CHAHC 146
RESULT 21
ID LEM2 RAT STANDARD; PRT; 549 AA.
AC P98105;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE OR ELAM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1].
RP SEQUENCE FROM N.A.

RC TISSUE=Lung;
RA Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SELECTIN/LECAM FAMILY.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L25527; AAA41113.1; -
CC HSP; P16581; 1KJA.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001304; Lectin C.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00059; lectin c; 1.
CC Pfam; PF00084; sushi; 5.
CC SMART; SM00032; CCP; 5.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00181; EGF; 1.
CC PROSITE; PS00022; EGF 1; 1.
CC PROSITE; PS01186; EGF 2; 1.
CC PROSITE; PS00615; C-TYPE LECTIN 1; 1.
CC PROSITE; PS00041; C-TYPE LECTIN 2; 1.
CC Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
CC Selectin; Signal; Sushi; Repeat.
CC SIGNAL 1 21
CC CHAIN 22 549
CC DOMAIN 22 494
CC EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 495 516
CC POTENTIAL.
CC DOMAIN 517 549
CC CYTOPLASMIC (POTENTIAL).
CC DOMAIN 38 138
CC C-TYPE LECTIN (SHORT FORM).
CC DOMAIN 139 175
CC EGF-LIKE.
CC DOMAIN 179 239
CC SUSHI 1.
CC DOMAIN 242 301
CC SUSHI 2.
CC DOMAIN 304 364
CC SUSHI 3.
CC DOMAIN 367 427
CC SUSHI 4.
CC DOMAIN 430 486
CC SUSHI 5.
CC BY SIMILARITY.
CC DISULFID 40 138
CC BY SIMILARITY.
CC DISULFID 111 130
CC BY SIMILARITY.
CC DISULFID 143 154
CC BY SIMILARITY.
CC DISULFID 148 163
CC BY SIMILARITY.
CC DISULFID 165 174
CC BY SIMILARITY.
CC DISULFID 180 225
CC BY SIMILARITY.
CC DISULFID 210 238
CC BY SIMILARITY.
CC DISULFID 243 287
CC BY SIMILARITY.
CC DISULFID 273 300
CC BY SIMILARITY.
CC DISULFID 305 350
CC BY SIMILARITY.
CC DISULFID 336 363
CC BY SIMILARITY.
CC DISULFID 368 413
CC BY SIMILARITY.
CC DISULFID 399 426
CC BY SIMILARITY.
CC DISULFID 431 472
CC BY SIMILARITY.
CC DISULFID 458 485
CC BY SIMILARITY.
CC CARBOHYD 25 25
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 60 60
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 145 145
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 192 192
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 203 203
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 266 266
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 313 313
CC N-LINKED (GLCNAC. . .) (POTENTIAL).

```
CC EMBL; X65700; CAA46622.1; -
DR EMBL; X65701; CAA46623.1; ALT_INIT.
DR EMBL; X40046; AAB22459.2; -
DR EMBL; M29279; AAA34089.1; -
DR EMBL; A16780; CAA01305.1; -
DR EMBL; X61679; CAA43854.1; -
DR HSP; P25871; LAUN.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; Thaumatin; 1.
DR PRINTS; PR00347; THAUMATIN.
DR ProDom; PD001321; Thaumatin; 1.
DR SMART; SM00205; THN; 1.
DR PROSITE; PS00316; THAUMATIN; 1.
KW Plant defense; Pathogenesis-related protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 246
FT DISULFID 30 225
FT DISULFID 72 82
FT DISULFID 87 93
FT DISULFID 141 213
FT DISULFID 146 196
FT DISULFID 154 164
FT DISULFID 168 177
FT DISULFID 178 183
FT CONFLICT 62 62
FT CONFLICT 122 122
FT CONFLICT 129 129
FT CONFLICT 132 132
FT CONFLICT 144 144
FT CONFLICT 149 149
FT CONFLICT 151 155
FT CONFLICT 156 156
FT CONFLICT 181 181
SQ SEQUENCE 246 AA; 26681 MW; 3CDBA991ACA2B0B1 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 246;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 4; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

Qy 1 CHAVDC 6
Db 141 CHAHC 146

RESULT 19
NP24 LYCES STANDARD; PRT; 247 AA.
AC P12670;
DT 01-OCT-1999 (Rel. 12, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NP24 protein precursor (Pathogenesis-related protein PR P23) (Salt-induced protein).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Aileas Craig;
RA Martin G.B., Jia Y.;
RT "Rapid transcript accumulation of pathogenesis-related genes during an incompatible interaction in bacterial speck disease resistant tomato plants."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 9-247 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. VFNT Cherry;
RA King G.J., Turner V.A., Husey C.E. Jr., Wurtele E.S., Lee S.M.;
RT "Isolation and characterization of a tomato cDNA clone which codes for a salt-induced protein."
RT a salt-induced protein.";
```

```
RL Plant Mol. Biol. 10:401-412(1988).
RN [3]
RP PARTIAL SEQUENCE.
RC STRAIN=cv. Rutgers; TISSUE=Leaf;
RX MEDLINE=91316233; PubMed=1859873;
RA Rodrigo I., Vera P., Frank R., Conejero V.;
RT "Identification of the viroid-induced tomato pathogenesis-related (PR) protein P23 as the thaumatin-like tomato protein NP24 associated with osmotic stress."
RL Plant Mol. Biol. 16:931-934(1991).
CC -1- SUBCELLULAR LOCATION: CYTOPLASM OR SOLUBLE FRACTIONS OF CYTOPLASMIC ORGANELLES, EXCEPT MITOCHONDRIA AND PLASTIDS.
CC -1- INDUCTION: BY SALT STRESS OR BY VIROIDS.
CC -1- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; AF093743; AAC64171.1; -
DR EMBL; M21346; AAA34175.1; -
DR PIR; JA0177; JA0177.
DR PIR; S07406; S07406.
DR PIR; S16264; S16264.
DR HSP; P25871; LAUN.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; Thaumatin; 1.
DR PRINTS; PR00347; THAUMATIN.
DR ProDom; PD001321; Thaumatin; 1.
DR SMART; SM00205; THN; 1.
DR PROSITE; PS00316; THAUMATIN; 1.
KW Plant defense; Pathogenesis-related protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 247
FT DISULFID 30 225
FT DISULFID 72 82
FT DISULFID 87 93
FT DISULFID 141 213
FT DISULFID 146 196
FT DISULFID 154 164
FT DISULFID 168 177
FT DISULFID 178 183
FT VARIANT 24 24
FT VARIANT I -> F (IN CV. RUTGERS)..
SQ SEQUENCE 247 AA; 26646 MW; 942E56E2531031B4 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 247;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CHAVDC 6
Db 141 CHAHC 146

RESULT 20
OS81 SOLCO STANDARD; PRT; 247 AA.
ID OS81 SOLCO
AC P50702;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Osmotin-like protein OSM181 precursor (PA81).
OS Solanum commersonii (Commerçon's wild potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4109;
RN [1]
RP SEQUENCE FROM N.A.
```

FT DISULFID 22 217 BY SIMILARITY.
 FT DISULFID 64 74 BY SIMILARITY.
 FT DISULFID 79 85 BY SIMILARITY.
 FT DISULFID 133 205 BY SIMILARITY.
 FT DISULFID 138 188 BY SIMILARITY.
 FT DISULFID 146 156 BY SIMILARITY.
 FT DISULFID 160 169 BY SIMILARITY.
 FT DISULFID 170 175 BY SIMILARITY.
 FT CONFLICT 69 69 D -> A (IN REF. 2).
 SQ SEQUENCE 238 AA; 25840 MW; 1BAE7CD4ABE4419 CRC64;
 Query Match 80.0%; Score 32; DB 1; Length 238;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 133 CHAHC 138

RESULT 17
 OS13_SOLCO STANDARD; PRT; 246 AA.
 ID OS13_SOLCO STANDARD; PRT; 246 AA.
 AC P50701;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE Osmotin-like protein OSML13 precursor (PA13).
 OS Solanum commersonii (Commerson's wild potato).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4109;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93192535; PubMed=8448373;
 RA Zhu B., Chen T.H.H., Li P.H.;
 RT "Activation of two osmotin-like protein genes by abiotic stimuli and
 fungal pathogen in transgenic potato plants.";
 RL Plant Physiol. 108:929-937(1995).
 CC -1- INDUCTION: BY ABSICISIC ACID (ABA), SALT, SALICYLIC ACID, WOUNDING,
 AND FUNGAL INFECTION.
 CC -1- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; X67121; CAA47601.1; --
 DR EMBL; X72928; CAA51432.1; --
 DR HSSP; P25871; IAUH.
 DR InterPro; IPR001938; Thaumatin.
 DR Pfam; PF00314; Thaumatin; 1.
 DR PRINTS; PR00347; THAUMATIN.
 DR ProDom; PD001321; Thaumatin; 1.
 DR SMART; SM00205; THN; 1.
 DR PROSITE; PS003316; THAUMATIN; 1.
 KW Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 246 OSMOTIN-LIKE PROTEIN OSML13.
 FT DISULFID 30 225 BY SIMILARITY.
 FT DISULFID 72 82 BY SIMILARITY.

FT DISULFID 87 93 BY SIMILARITY.
 FT DISULFID 141 213 BY SIMILARITY.
 FT DISULFID 146 196 BY SIMILARITY.
 FT DISULFID 154 164 BY SIMILARITY.
 FT DISULFID 168 177 BY SIMILARITY.
 FT DISULFID 178 183 BY SIMILARITY.
 SQ SEQUENCE 246 AA; 26654 MW; 5A9C5A5A1910732 CRC64;
 Query Match 80.0%; Score 32; DB 1; Length 246;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6
 DB 141 CHAHC 146

RESULT 18
 OSMO_TOBAC STANDARD; PRT; 246 AA.
 ID OSMO_TOBAC STANDARD; PRT; 246 AA.
 AC P14170;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Osmotin precursor.
 DE AP24.
 GN Nicotiana tabacum (Common tobacco).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Samsun NN; TISSUE=Leaf;
 RX MEDLINE=93192519; PubMed=8448358;
 RA Melchers L.S.; Sela-Buurlage M.B.; Vloemans S.A.; Woloshuk C.P.;
 RA van Roekel J.S.-C.; Pen J.; van den Elzen P.J.M.; Cornelissen B.J.C.;
 RT "Extracellular targeting of the vacuolar tobacco protein AP24,
 chitinase and beta-1,3-glucanase in transgenic plants.";
 RL Plant Mol. Biol. 21:583-593(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92329718; PubMed=1385735;
 RA Nelson D.E.; Raghothama K.G.; Singh N.K.; Hasegawa P.M.; Bressan R.A.;
 RT "Analysis of structure and transcriptional activation of an osmotin
 gene.";
 RL Plant Mol. Biol. 19:577-588(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Wisconsin 38;
 RA Singh N.K.; Nelson D.E.; Kuhn D.; Hasegawa P.M.; Bressan R.A.;
 RT "Molecular cloning of osmotin and regulation of its expression by ABA
 and adaptation to low water potential.";
 RL Plant Physiol. 90:1096-1101(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. White Burley;
 RX MEDLINE=92163032; PubMed=1536937;
 RA Kumar V.; Spencer M.B.;
 RT "Nucleotide sequence of an osmotin cDNA from the Nicotiana tabacum
 cv. white burley generated by the polymerase chain reaction.";
 RL Plant Mol. Biol. 18:621-622(1992).
 CC -1- SUBCELLULAR LOCATION: VACUOLAR INCLUSION BODIES.
 CC -1- INDUCTION: BY SALT STRESS AND BY ABSICISIC ACID (ABA).
 CC -1- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

RESULT 15

Y296 HUMAN
ID Y296 HUMAN STANDARD; PRT; 1829 AA.
AC O15015;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical zinc finger protein KIA0296.
GN KIA0296
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB002294; BAA20756.1; -
DR InterPro: IPR000822; Znf.C2H2.
DR Pfam: PF00096; zf.C2H2; 31.
DR PRINTS; PRO0048; ZINC.FINGER.
DR SMART; SM00355; Znf.C2H2; 30.
DR PROSITE; PS00028; ZINC.FINGER.C2H2.1; 30.
DR PROSITE; PS0157; ZINC.FINGER.C2H2.2; 29.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Zinc-finger; Metal-binding; Nuclear protein; Repeat.
FT DOMAIN 8 1783
FT ZN_FING 8 31
FT ZN_FING 48 70
FT ZN_FING 75 97
FT ZN_FING 239 261
FT ZN_FING 266 288
FT ZN_FING 294 316
FT ZN_FING 374 396
FT ZN_FING 401 424
FT ZN_FING 465 487
FT ZN_FING 492 514
FT ZN_FING 575 597
FT ZN_FING 617 639
FT ZN_FING 644 666
FT ZN_FING 821 843
FT ZN_FING 848 870
FT ZN_FING 881 904
FT ZN_FING 958 980
FT ZN_FING 1052 1074
FT ZN_FING 1079 1101
FT ZN_FING 1203 1225
FT ZN_FING 1230 1252
FT ZN_FING 1258 1280
FT ZN_FING 1299 1321
FT ZN_FING 1326 1348
FT ZN_FING 1364 1386
FT ZN_FING 1364 1386

FT ZN_FING 1557 1579 C2H2-TYPE.
FT ZN_FING 1585 1607 C2H2-TYPE.
FT ZN_FING 1677 1699 C2H2-TYPE.
FT ZN_FING 1704 1726 C2H2-TYPE.
FT ZN_FING 1732 1754 C2H2-TYPE.
FT ZN_FING 1761 1783 C2H2-TYPE.
SQ SEQUENCE 1829 AA; 200760 MW; B20C482B38684895 CRC64;
Query Match 82.5%; Score 33; DB 1; Length 1829;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 CHAVDC 6
Db 1498 CHAGDC 1503
RESULT 16
TPM1 LYCES
ID TPM1 LYCES STANDARD; PRT; 238 AA.
AC Q01591;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osmotin-like protein TPM-1 precursor (PR P23) (Fragment).
GN TPM-1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Rio Grande; TISSUE=Leaf;
RX MEDLINE=93099273; PubMed=1463856;
RA Ruiz-Medrano R., Jlmenez-Moralla B., Herrera-Estrella L.,
RA Rivera-Bustamante R.F.;
RT "Nucleotide sequence of an osmotin-like cDNA induced in tomato during
RT viroid infection.";
RL Plant Mol. Biol. 20:1199-1202(1992).
RN [2]
RP SEQUENCE OF 6-238 FROM N.A.
RC STRAIN=cv. Rutgers; TISSUE=Leaf;
RX MEDLINE=94105335; PubMed=8278538;
RA Rodrigo I., Vera P., Tornero P., Hernandez-Yago J., Conejero V.;
RT "cDNA cloning of viroid-induced tomato pathogenesis-related protein
RT P23. Characterization as a vacuolar antifungal factor.";
RL Plant Physiol. 102:939-945(1993).
CC -!- FUNCTION: ANTIFUNGAL PROTEIN.
CC -!- INDUCTION: BY VIROID INFECTION.
CC -!- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X66416; CAA47047.1; -
DR EMBL; X70787; CAA50059.1; -
DR PIR; S28001; S28001.
DR HSSP; P25871; LAUN.
DR InterPro: IPR001938; Thaumatin.
DR Pfam; PF00314; Thaumatin; 1.
DR ProDom; PD001321; Thaumatin; 1.
DR SMART; SM00205; THN; 1.
DR PROSITE; PS00316; THAUMATIN; 1.
KW Plant defense; Pathogenesis-related protein; Signal; Fungicide.
FT NON_TER 1
FT SIGNAL <1 13
FT CHAIN 14 238
FT OSMOTIN-LIKE PROTEIN TPM-1.

RP SEQUENCE FROM N.A.
 RC STRAIN=SRRC 143 / ATCC 56775, and SRRC 2043 / ATCC 62882;
 RX MEDLINE=99054922; PubMed=9835571;
 RA Yu J., Chang P.-K., Ehrlich K.C., Cary J.W., Montalbano B., Dyer J.M.,
 RA Bhatnagar D., Cleveland T.E.;
 RA "Characterization of the critical amino acids of an Aspergillus
 RT parasiticus Cytochrome P-450 monooxygenase encoded by orda that is
 RT involved in the biosynthesis of aflatoxins B1, G1, B2, and G2.";
 RT Appl. Environ. Microbiol. 64:4834-4841(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 163 / NRRL 5862 / SU-1;
 RX MEDLINE=20312409; PubMed=10855719;
 RA Yu J., Chang P.-K., Bhatnagar D., Cleveland T.E.;
 RT "Genes encoding cytochrome P450 and monooxygenase enzymes define one
 RT end of the aflatoxin pathway gene cluster in Aspergillus
 RT parasiticus.";
 RT Appl. Microbiol. Biotechnol. 53:583-590(2000).
 RL -!- FUNCTION: CONVERTS O-METHYLSTERIGMATOCYSTIN (OMST) TO AFLATOXINS
 CC AFLATOXINS B2 OR G2.
 CC -!- PATHWAY: Aflatoxin biosynthesis.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF017151; AAC95058.1; -
 CC EMBL; AF169016; AAF26278.1; -
 CC InterPro; IPR001128; Cytochrome_P450.
 CC Pfam; PF00067; P450; 2.
 CC PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Heme.
 FT BINDING 440 440 HEME (BY SIMILARITY).
 FT VARIANT 143 143 A -> S (IN STRAIN SRRC 2043).
 FT VARIANT 400 400 H -> L (IN STRAIN SRRC 2043; LOSS OF
 FT ACTIVITY).
 FT VARIANT 528 528 I -> Y (IN STRAIN SRRC 2043).
 SQ SEQUENCE 528 AA; 60188 MW; AC32F9425D27F328 CRC64;

Query Match 82.5%; Score 33; DB 1; Length 528;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6

Db 455 CHAISC 460

RESULT 14

HIG DROME STANDARD; PRT; 958 AA.
 AC Q09101;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Locomotion-related protein Hikaru genki precursor.
 GN HIG
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 P TISSUE=Head;
 MEDLINE=93213498; PubMed=8461133;
 'Shino M., Matsuzaki F., Nabeshima Y.-I., Hama C.;

RT "Hikaru genki, a CNS-specific gene identified by abnormal locomotion
 RT in Drosophila, encodes a novel type of protein.";
 RL Neuron 10:395-407(1993).
 CC -!- FUNCTION: HAS A ROLE IN THE DEVELOPMENT OF CNS FUNCTIONS INVOLVED
 CC IN LOCOMOTOR ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1, 2, 3 (SHOWN HERE) AND 4; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN PCC NEURONS AND NEUROBLASTS
 CC IN THE PROCEPHALIC NEUROGENIC REGION IN THE CENTRAL NERVOUS
 CC SYSTEM.
 CC -!- DEVELOPMENTAL STAGE: MOST ABUNDANT DURING AND/OR AFTER NEURONAL
 CC DIFFERENTIATION AND DURING CELL SPECIFICATION OR AXOGENESIS.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D13884; BAA02984.1; -
 CC EMBL; D13885; BAA02985.1; -
 CC EMBL; D13886; BAA02986.1; -
 CC EMBL; D13887; BAA02987.1; -
 CC HSP; P10998; 1VVD.
 CC FlyBase; FBgn0010114; hig.
 CC InterPro; IPR003599; Ig.
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR000436; Sushi_SCR_CCP.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF00084; sushi; 5.
 CC SMART; SM00032; CCP; 5.
 CC SMART; SM00409; Ig; 1.
 KW Glycoprotein; Alternative splicing; Immunoglobulin domain; Repeat;
 FT SIGNAL 1 31
 FT CHAIN 32 958
 FT DOMAIN 630 709
 FT DOMAIN 713 769
 FT DOMAIN 772 828
 FT DOMAIN 831 891
 FT DOMAIN 893 952
 FT SITE 318 320
 FT DISULFID 714 755
 FT DISULFID 741 768
 FT DISULFID 773 814
 FT DISULFID 800 827
 FT DISULFID 832 877
 FT DISULFID 863 890
 FT DISULFID 894 939
 FT DISULFID 922 952
 FT CARBOHYD 376 376
 FT CARBOHYD 525 525
 FT CARBOHYD 605 605
 FT CARBOHYD 620 620
 FT CARBOHYD 752 752
 FT CARBOHYD 789 789
 FT CARBOHYD 529 553
 FT VARSPPLIC 892 958
 SQ SEQUENCE 958 AA; 107027 MW; 4161258E95ABC764 CRC64;

Query Match 82.5%; Score 33; DB 1; Length 958;
 Best Local Similarity 66.7%; Pred. No. 70;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CHAVDC 6

Db 768 CHAIOC 773